

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 21:00:05 ; Search time 16233 Seconds
(without alignments)
7008.899 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625

Sequence: 1 atgagctgagctgctctcac.....tcaaaaaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenDbml.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2624	100.0	2625	8 U73588	U73588 Gossypium h
2	1689.6	64.4	2661	8 AB022092	AB022092 Citrus un
3	1615.6	61.5	2652	8 VIRUS81	D10266 Vigna radia
4	1602.6	61.1	2842	8 AF030231	AF030231 Glycine m
5	1577.2	60.1	2738	8 AF315375	AF315375 Phaseolus
6	1559	59.4	2728	8 MTR11943	MTR11943 Medicago
7	1556	59.3	2652	8 PSA012080	PSA012080 Pium sat
8	1556	59.3	2652	8 AF049487	AF049487 Medicago
9	1555.6	59.3	2770	8 AB018561	AB018561 Citrullus
10	1548.6	59.0	2749	8 AF079851	AF079851 Pium sat
11	1539	58.6	2647	8 VFSUDCS	X69773 V.faba mRNA
12	1539	58.6	2665	8 VFAUDPGFTA	M97551 Vicia faba
13	1499.4	57.1	2906	6 BD235993	BD235993 Materials
14	1497.8	57.1	2913	6 BD235993	BD235993 Materials
15	1497.8	57.1	3103	6 BD262165	BD262165 Compoviti
16	1497.8	57.1	3103	6 AR360950	AR360950 Sequence
17	1486.6	56.6	2427	6 AX506925	AX506925 Sequence
18	1486.6	56.6	2427	6 AX651931	AX651931 Sequence
19	1477.6	56.3	2783	8 AGSUS1	X92378 A. glutinosa
20	1475.6	56.2	2757	8 PSA311496	PSA311496 Pium sat
21	1454	55.4	2725	8 TOMSSF	L19762 Lycopersico
22	1452.4	55.3	2711	8 POTSSFN	M18745 Potato sucr
23	1447.8	55.2	2708	8 LES011319	AJ011319 Lycopersi
24	1447.6	55.1	2701	8 AY205084	AY205084 Solanum t
25	1442.8	55.0	2429	8 STUS57575	AY57575 Solanum t
26	1405.4	53.5	2838	8 AY457173	AY457173 Beta vulg
27	1396.4	53.2	2554	8 CRSUCSYN	X82504 C.rudrum mk
28	1380.8	52.6	2866	8 DCRNAS	X75332 D.carota (N
29	1316	50.1	2717	8 AF263384	AF263384 Saccharum
30	1315.4	50.1	2910	8 AF412037	AF412037 Bambusa o
31	1313.6	50.0	2771	8 AK100306	AK100306 Oryza sat
32	1310.2	49.9	2690	8 AF412038	AF412038 Bambusa o
33	1307	49.8	2774	8 AF412039	AF412039 Bambusa o
34	1301.8	49.6	2890	8 AF412036	AF412036 Bambusa o
35	1300.6	49.5	2494	6 E55472	E55472 Method for
36	1300.6	49.5	3389	8 AK100546	AK100546 Oryza sat
37	1299	49.5	2702	8 AK098923	AK098923 Oryza sat
38	1287.4	49.4	2627	8 OSSUPHSY	Z15078 O.sativa mr
39	1286.2	49.4	2885	8 TGSUDCSY21	X96939 T.gesnerian
40	1285.8	49.4	2781	8 AK072074	AK072074 Oryza sat
41	1287.6	49.1	2450	6 AX654790	AX654790 Sequence
42	1287.6	49.1	2451	6 AX755420	AX755420 Sequence
43	1287.6	49.1	2451	6 AX755773	AX755773 Sequence
44	1287.6	49.1	2910	8 AK100334	AK100334 Oryza sat
45	1285.4	49.0	2759	8 TGSUDCSYNI	X96938 T.gesnerian

ALIGNMENTS

RESULT 1
LOCUS U73588 2625 bp mRNA linear PLN 04-MAY-1999
DEFINITION Gossypium hirsutum sucrose synthase mRNA, complete cds.
ACCESSION U73588
VERSION U73588.2 GI:4733945
KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
AUTHORS Perez-Grau, L. and Delmer, D.
TITLE Direct Submission

JOURNAL	Submitted (07-OCT-1996) Calgene, Inc., 1920 Fifth Street, Davis, CA 95616, USA
COMMENT	On May 4, 1999 this sequence version replaced gi:4098126.
FEATURES	Location/Qualifiers
SOURCE	1..2625

```
source
1.2635
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/strail="Acala SJ-2"
/db_xref="taxon:3635"
/tissue_type="cotton fiber"
1.2421
```

```

/codon start=1
/product="sucrose synthase"
/protrein_id="FAD28641.1"
/db_xref="GI:4733946"
/translation="MWRRLTRVSHSLRELDITLARNETLALSLTBGGKIGILQH
HOLLVEEALVEENRRKILANGAEVDESVKASQELI VLRPWVALVRPGEWETIRV
HALLVBEELTAEENRKLKEELVDESSNCPVLEIDFEEPSNFRPRLTSKISGVGEV
NRLHSLKFLPDKSSMHPLEFLPLVTHCHKGNMMLNRIQINMLAOHLRAAEVGLT
PEPTCAEAFENRPOELGLEKGMWDTRAEVLEMLQLDLLEADPCLEKFAEGLIPV
FNVVITPLPGYFQODLVKAGVPTDGGQVYLIDVDRALENMLRIKQOGLNITPRLI
ITFLLEDAVGTGQADNRKLVKYGESHSDILVPRTEGIVKRMISREKWEPLAET
EDVAKHISKEMLDTPDILFAMSDGNVSLALHAKGVQCTLAHLERKYPDSOIV
MKCLEKHSKSCPTDILFAMSDGNVSLALHAKGVQCTLAHLERKYPDSOIV
VHGLIDVPRKPKVSPGADMLVPRTEKREKTHHPREIDILVYKVNESHLCV
NDNRKRIILTPMLDYKVKITGLVGVENCKPKIRBLANLVVGGDRKSEKLEBEK
MKKNEFLIDKTNLNGORWASSQNNRIRNRELRIITDITGAFQVPLAETGLTVE
AATCGATPACNGPHEILVHSGSDNITPYGDAQDLVLPFECKRQPSMDKI
SQGGLRIEKKETWKLVSERLLTLGVYGVFWKIVSNLREESRRYLETMPALYKRLA
EVSFLAAE"
2608..2625

```

polya_site
ORIGIN

Query Match	100.0%	Score 2524	DB 8	Length 2625
Best Local Similarly	100.0%	Pred. No. 0		
Matches 2625; Conservative	0	Mismatches	0	Gaps 0

Qy	1	ATGCTAGACGCTGCTCTCACTCGGCTCCAGACGCTCCGCTGAGCGCTTGGATGAGACCCCTT	60
Db	1	ATGGCTAGCGCTGCTCTCACTCGGCTCCAGACGCTCCGCTGAGCGCTTGGATGAGACCCCTT	60
Qy	61	CTTGCTCCAGAGAAAGAGATTTTGGCTCTGCTCTCAAGATCGAGGGCAAGAGAAAGAA	120
Db	61	CTTGCTCCAGAGAAAGAGATTTTGGCTCTCTCAAGATCGAGGGCAAGAGAAAGAA	120
Qy	121	ATTTCGCAACCATCAATTAATTCYAGATTTGAAGTATCCCTGAAAGAAACAGAAAG	180
Db	121	ATTTCGCAACCATCAATTAATTCYAGATTTGAAGTATCCCTGAAAGAAACAGAAAG	180
Qy	181	AAGCTCGCTAATGAGTCAATTTTTCAGATATTTGAAGCTAGTCAGAGACGATCGTGTG	240
Db	181	AAGCTCGCTAATGAGTCAATTTTTCAGATATTTGAAGCTAGTCAGAGACGATCGTGTG	240
Qy	241	CCTTCATAGGGTTGCACTTGTGCTGCTGCTCCAGAGCGCTGATGTTTGGAGTACATTAGAGT	300
Db	241	CCTTCATAGGGTTGCACTTGTGCTGCTGCTCCAGAGCGCTGATGTTTGGAGTACATTAGAGT	300
Qy	301	AATGTTCAAGCGCTTGTGTGTGAGGAACTCACTGTTGCTGAGTATCTCACTTCAAGGAA	360
Db	301	AATGTTCAAGCGCTTGTGTGTGAGGAACTCACTGTTGCTGAGTATCTCACTTCAAGGAA	360
Qy	361	GAGCTTGTGATGGAAGTTCAAAATGAAACTTTGTTTGGAAATGGAATTTTGAAGCCCTTC	420
Db	361	GAGCTTGTGATGGAAGTTCAAAATGAAACTTTGTTTGGAAATGGAATTTTGAAGCCCTTC	420
Qy	421	AATCATCATTCGCCCGCCCAACTCTTTCAAAAATTCATTGGTAAATGAGTGAATTCCTA	480
Db	421	AATCATCATTCGCCCGCCCAACTCTTTCAAAAATTCATTGGTAAATGAGTGAATTCCTA	480
Qy	481	AATGTCACCTTTTGCCCAAAATGTTTCATGACAGAGAGCATGCACTCTTGTCTGAA	540
Db	481	AATGTCACCTTTTGCCCAAAATGTTTCATGACAGAGAGCATGCACTCTTGTCTGAA	540

QY	541	TTCTCTGAGATCCATTGTGCACAGGGCGAACAATGATGTTGAAATGACAAATTCAGAAC	600
Db	541	TTCTCTGAGATCCATTGTGCACAGGGCGAACAATGATGTTGAAATGACAAATTCAGAAC	600
QY	601	TTGAATGCTCTTCAACATGTTTTTGAAGAAAGCAGAGAGATCTTGGTACCTCTCTCT	660
Db	601	TTGAATGCTCTTCAACATGTTTTTGAAGAAAGCAGAGAGATCTTGGTACCTCTCTCT	660
QY	661	GAGACACCATGTGCGGAATTCGAACACCGGTTCCAGGAAATCGGTTTTGAAAGAAGTTGG	720
Db	661	GAGACACCATGTGCGGAATTCGAACACCGGTTCCAGGAAATCGGTTTTGAAAGAAGTTGG	720
QY	721	GGTACACCCGACAGAACCGCTGCTGAGATGATCCAATCTCTTTGGATCTTCTTGAGCA	780
Db	721	GGTACACCCGACAGAACCGCTGCTGAGATGATCCAATCTCTTTGGATCTTCTTGAGCA	780
QY	781	ACTGATCCTTGCACCCCTTGGAAGATGTCCTTGGGAGAAATCCCATGSGTTCAAATGTTGTG	840
Db	781	ACTGATCCTTGCACCCCTTGGAAGATGTCCTTGGGAGAAATCCCATGSGTTCAAATGTTGTG	840
QY	841	ATTCTCACTCCCCACGGATACTTCGCTCAAGACAAATGTTTTGGGATATCCCGACACCGGT	900
Db	841	ATTCTCACTCCCCACGGATACTTCGCTCAAGACAAATGTTTTGGGATATCCCGACACCGGT	900
QY	901	GGGCAAGTGTGTTACATCTTGGATCAAGTCGAGCTTTGGAGAAATGAGATGCTCTCCGT	960
Db	901	GGGCAAGTGTGTTACATCTTGGATCAAGTCGAGCTTTGGAGAAATGAGATGCTCTCCGT	960
QY	961	ATTAAAGAACAAAGACATCAATCAACCCCTCGAATCCCATATTACTAGACTTCTCCCT	1020
Db	961	ATTAAAGAACAAAGACATCAATCAACCCCTCGAATCCCATATTACTAGACTTCTCCCT	1020
QY	1021	GATGCTGTCCGAAACAAATGCGGTCAACGACTTGAGAAAGTATACGGAACAGACACTCG	1080
Db	1021	GATGCTGTCCGAAACAAATGCGGTCAACGACTTGAGAAAGTATACGGAACAGACACTCG	1080
QY	1081	GATATCTCTTGAGTACCTCTTCAGAACAAAGGGAATGTTGCGAAATAGATCTTCAAGA	1140
Db	1081	GATATCTCTTGAGTACCTCTTCAGAACAAAGGGAATGTTGCGAAATAGATCTTCAAGA	1140
QY	1141	TTTGAAGAAAGCTGGCCATATCTTGGAACCTTACACAGAGATGTTGTCATGAAATCTCC	1200
Db	1141	TTTGAAGAAAGCTGGCCATATCTTGGAACCTTACACAGAGATGTTGTCATGAAATCTCC	1200
QY	1201	AAAGAGTTGACGGCACGCGCAGATCTGTATCATCGGAACAAACAGCGACATATGCTC	1260
Db	1201	AAAGAGTTGACGGCACGCGCAGATCTGTATCATCGGAACAAACAGCGACATATGCTC	1260
QY	1261	GCTCTCTTGTCTCGACATAAATTAGTGTCAACACGTGCACATCGCCCATGCTTTGGAG	1320
Db	1261	GCTCTCTTGTCTCGACATAAATTAGTGTCAACACGTGCACATCGCCCATGCTTTGGAG	1320
QY	1321	AAGCAAAATATCCAGATTCAAGATATCTATTGGAAGAAAGCTTGGAACAAATACATTTCC	1380
Db	1321	AAGCAAAATATCCAGATTCAAGATATCTATTGGAAGAAAGCTTGGAACAAATACATTTCC	1380
QY	1381	TCTTGCCAAATTACAGCTGATCTTTTGGCAATGAACATACAGATTTCACTATCACAGT	1440
Db	1381	TCTTGCCAAATTACAGCTGATCTTTTGGCAATGAACATACAGATTTCACTATCACAGT	1440
QY	1441	ACTTTCAGAGAAATTGACAGAAAGCAACATGTTGTCAAATCGAAGCCACTGCT	1500
Db	1441	ACTTTCAGAGAAATTGACAGAAAGCAACATGTTGTCAAATCGAAGCCACTGCT	1500
QY	1501	TTCACTCTTCTGATCTCTACCGGTGTGTACATGATATCGATGTGTTGATATCCAAATTC	1560
Db	1501	TTCACTCTTCTGATCTCTACCGGTGTGTACATGATATCGATGTGTTGATATCCAAATTC	1560
QY	1561	AAACATGTTTCCCTGGTGTGATATGAGATATACTCCCTTACACCGAAGAGAAAGCGG	1620
Db	1561	AAACATGTTTCCCTGGTGTGATATGAGATATACTCCCTTACACCGAAGAGAAAGCGG	1620
QY	1621	AGGTGGAAGATTTCCATCTCTGAGATGAGAACCTTTCTTTACACAAAGTTGAGAAATGA	1680

```

Db      1621 AGGTGAAGCATTTTCATCTGAGATCGAAGACCTTTTACACCAAGTTGAGATGAA 1680
Qy      1681 GAACTATATGTGTCCTAATGACCGCAACAAGCCATTTCTGTTACAAATGCCAAGGCTT 1740
Db      1681 GAACTATATGTGTCCTAATGACCGCAACAAGCCATTTCTGTTACAAATGCCAAGGCTT 1740
Qy      1741 GATCGTGTCAAGAACTTAACCGGACTCGTCGAGTGGTGGCGGAGAACCCAAAGTTGGCT 1800
Db      1741 GATCGTGTCAAGAACTTAACCGGACTCGTCGAGTGGTGGCGGAGAACCCAAAGTTGGCT 1800
Qy      1801 GAGTTGGCTTAACCTGTAAGTTGTAGTGTGTGTAGCGCAAGAAAGAACTTAAAGATTGGAA 1860
Db      1801 GAGTTGGCTTAACCTGTAAGTTGTAGTGTGTGTAGCGCAAGAAAGAACTTAAAGATTGGAA 1860
Qy      1861 GAGAAAGCTGTAATGAAAGAAATGTTGAGCTGATCGAAGTACAACTTGAACGGCCAA 1920
Db      1861 GAGAAAGCTGTAATGAAAGAAATGTTGAGCTGATCGAAGTACAACTTGAACGGCCAA 1920
Qy      1921 TTCAGATGATATCATCTCAATGAAAGAAATCCGAAAGTTGAACCTTACCGATACATT 1980
Db      1921 TTCAGATGATATCATCTCAATGAAAGAAATCCGAAAGTTGAACCTTACCGATACATT 1980
Qy      1981 TCGCAACAGAAAGGTGCTTTGTACAGGCTGCTGATTTGAAGCCTTTGATTTGACAGTT 2040
Db      1981 TCGCAACAGAAAGGTGCTTTGTACAGGCTGCTGATTTGAAGCCTTTGATTTGACAGTT 2040
Qy      2041 GTGAGGCGAATGACTTGGGCTTTGCGCAACTTGGCAACCTTGAACGGTGAACACCGGAG 2100
Db      2041 GTGAGGCGAATGACTTGGGCTTTGCGCAACTTGGCAACCTTGAACGGTGAACACCGGAG 2100
Qy      2101 ATTAATTCGCAAGGAAATCTGTTTCAACATGATCCCTTACATGATGATCAAGCTCT 2160
Db      2101 ATTAATTCGCAAGGAAATCTGTTTCAACATGATCCCTTACATGATGATCAAGCTCT 2160
Qy      2161 GACATACTCGTCGATTTCTTGAAGAAAGTGTAGAAAGATCCATCTCACTGGGATTAAGATC 2220
Db      2161 GACATACTCGTCGATTTCTTGAAGAAAGTGTAGAAAGATCCATCTCACTGGGATTAAGATC 2220
Qy      2221 TCCCAAGAGAGCTTGAAGCAATGAGAGAAAGTATACATGAGAGATTTACTCGGAGAGA 2280
Db      2221 TCCCAAGAGAGCTTGAAGCAATGAGAGAAAGTATACATGAGAGATTTACTCGGAGAGA 2280
Qy      2281 CTATTGACCTCGAAGAGAGTATGATTTCCGAGACATGTTTCCAACTTGAAGCGCGT 2340
Db      2281 CTATTGACCTCGAAGAGAGTATGATTTCCGAGACATGTTTCCAACTTGAAGCGCGT 2340
Qy      2341 GAGAGTCGTGCTTACCTTGAGATGTTTATGCTCTTAAAGTACCGTAACTGGCTGAATCA 2400
Db      2341 GAGAGTCGTGCTTACCTTGAGATGTTTATGCTCTTAAAGTACCGTAACTGGCTGAATCA 2400
Qy      2401 GTTCCATTGGCAGAGAGATTAATTGAACCTGTTAAATACATGAGGCGGCTTTTCTTGG 2460
Db      2401 GTTCCATTGGCAGAGAGATTAATTGAACCTGTTAAATACATGAGGCGGCTTTTCTTGG 2460
Qy      2461 AGAATAATATCTGTTTGTAAATTTCAATGAGAGAGTCCCTTGTATTTCAATCTGTCT 2520
Db      2461 AGAATAATATCTGTTTGTAAATTTCAATGAGAGAGTCCCTTGTATTTCAATCTGTCT 2520
Qy      2521 TTTCCCTTTCTTTTTCGCGCGCATTTGTTGAACATGGGGTGTGCGCCGCTCAATTC 2580
Db      2521 TTTCCCTTTCTTTTTCGCGCGCATTTGTTGAACATGGGGTGTGCGCCGCTCAATTC 2580
Qy      2581 AGTTAAATATGTGACTTTTGTTCAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCA 2625
Db      2581 AGTTAAATATGTGACTTTTGTTCAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCA 2625

```

RESULT 2
 AB022092 2661 bp mRNA linear PLN 25-JAN-2002
 LOCUS
 DEFINITION Citrus unshiu CItSUS1 mRNA for sucrose synthase, complete cds.
 ACCESSION AB022092

```

VERSION AB022092.1 GI:6682842
KEYWORDS sucrose synthase.
SOURCE Citrus unshiu
ORGANISM Citrus unshiu
REFERENCE 1
AUTHORS Komatsu,A., Moriguchi,T., Koyama,K., Omura,M. and Akihama,T.
TITLE Analysis of sucrose synthase genes in citrus suggests different
JOURNAL roles and phylogenetic relationships
MEDLINE J. Exp. Bot. 53 (366), 61-71 (2002)
PUBMED 11741042
REFERENCE 2 (bases 1 to 2661)
AUTHORS Komatsu,A.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1998) Akira Komatsu, National Institute of Crop
Science, 2-1-18, Tsukuba Science City, Ibaraki 305-8518, Japan
(B-mail:akomatsu@affrc.go.jp, Tel:81-298-38-8949, Fax:81-298-38-8949)
FEATURES
Source
1. .2661
/organism="Citrus unshiu"
/mol_type="mRNA"
/cultivar="Miyagawa-Wase"
/db_xref="taxon:55188"
/tissue_type="fruit"
/dev_stage="middle developmental stage of fruit juice
sacs"
1. .2661
/gene="CItSUS1"
24. .2441
/gene="CItSUS1"
/codon_start=1
/product="sucrose synthase"
/protein_id="BA88905.1"
/db_xref="GI:6682843"
/translation="MAERALTVHSLREIDETLSAHRBILALLSRIEGKGLION
HOLIAEPESISEERKHLTEGAREVLRATOEALVLAAPVLAAPVPRGVEYIRAVY
HALVEELVAEYLAHFEELVDGSSNPFVLEDFEPNAPFPPTLSKISGCEVFL
NRLHSAFLPHDKESMPLLEFLRVHCKGNMMLNDRIQINLSQHLRAEYLLTV
VPEPFEELALRFOEIGLERGMDPTARALEMIDLDLEAPPTLEFLGRIPMV
FNVLPHGFADDDVGVDPDGGOVVILDOVRALDEMLRIKOGGLDITPOIL
ITRLPRAVGTGCRLEKYCTGKSLILVPRTEGVRKXISREVPYLETYE
DNVAEIKELDGKPDILIGNYSDGNIVASLHAGLGTCTTAALEKTKPDSDIYR
KNLDKXHFSCQFTADLIANNHTDPIITSTFOELAGSKDVGQYESHAFLLPELYRV
VHGIDVDPKRNIVSPGADMSIYFPYEBKRLRSFPIBELLYSDVENKEHLGVK
DRNKPILEPTARLDVRKVLGLVYVYKNAKRLRLNVLVVGDRRKESKDLERQAEK
KRWYSLIDONKNGOFWISSOMNRVRNGLHYRICETKGAFYOPALYEAAGLVVVA
MTGSLPTPTNCKSPAEIIVNGKSGHIDPIHGEQAAELVDPEKCKADPSVYDKTS
DGLKRIEBKTKIYISORLUTLTVGVGFVHNSLDRLESRRYLEMFLVKRLAE
SVPLAVE"
2661
/gene="CItSUS1"
/note="10 a nucleotides"
ORIGIN
Query Match 64.4%; Score 1689.6; DB 8; Length 2661;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 485; Indels 3; Gaps 1;
polyA_site
1 ATGGCTGAGCGTGTCTTCATCTGCGTCAACAGTCTCCGTGAGCGTTTGATGAACCTT 60
24 ATGGCCGAAACGTGCTTTGACCCCGGTTCAAGTCTCCGCGAGCGTTTGATGAACCTT 83
61 CTGCTCAGAGAGAGATTTTGGCTTGGCTCTCAAGATCGAGGCGAAAGAAAAGA 120
84 TCTGCCACAGAGATGAATCTTGGCCCTTTGTGCAAGATGAGGCGCAAGAGA 143
121 ATTCTGCAACCAATCAATATTATTCTAGAGTTTGAAGCTATCCTTGAAGACAGAAG 180
144 ATCTTGCAAAACATCAATATTGCTGAGTTTGAATCAATTTCTGAGGAAAAAG 203

```

Qy	181	AAGTCGCTAATGSGTGCATTTTGTGAAATGATGAGGCTAGTCAGAGACGATCGTGTG	240
Db	204	CATCTAATCTAGGGTGCATTTGGTGAAGTCTTAAGGCCACCCAGAAACAAATGTGTG	263
Qy	241	CTTCCATGGGTGCATTTGCTGTTCGTCAAAGGCTGTGTGTGGAGTACATTAGAGTG	300
Db	264	GCTCCGTGGGTAAGCGCTGCGGTGCGTCMAAGGCCGGGTGTGGAGTACATCCGAGTGG	323
Qy	301	AATGTTCAAGCGCCCTGTGTTGTGAGAACTCACTGTTGTGAGATGTCACATTCAGAGA	360
Db	324	AATGTCCATGCGCTGTGTCGTGTGAGAACTCTGTATGCGAGTACCTGCATTCAGAGAG	383
Qy	361	GAGCTGTGTATGAGAAATTGAAATGAAACTTGTGTGTGAAATTTGATTTTGAAGCCCTTC	420
Db	384	GAACTGTTCATGAGAGGCTCCAAATGGCACTTGCTGCTTGAGTGGATTTTGAAGCAATTC	443
Qy	421	AATCATCATTTCCCGCGCCCACTCTTTCAAATTCATTTGTTAATGTGTGAGTTCTTA	480
Db	444	AATCCCTGTGTCTCTCCCAACTCTTTCMAAGTCTAATGGAATGTGTGAATTTCTCT	503
Qy	481	AATGTCACTTTTGAGAAATTTGTTCCATGACAAAGAGAGATGACCTTTGCTCGAA	540
Db	504	AACCGCACTTTTGTGGAAGCTCTTTCATGACAAAGAGAGATGATCACTGCTTGAA	563
Qy	541	TTCTCTGAGATGCATTTGTCACAAGGCGCAGACATGATGTTGTAATGACAAATTCAGAC	600
Db	564	TTTCTCCGCGTCCACTCCCAAGGGCAAAACATGATGCTGAAATGACAAATTCAGAAC	623
Qy	601	TTGAAATGCTCTTCAACATGTTTTGAGAAAGCAGAGAGATTCCTTGATCCCTACTCTCT	660
Db	624	TTGAAATTCCTTCAACATGTTCTGAGAGAGCAGAGAGATTCGACCAAGTGTATCA	683
Qy	661	GAGACACCATGTGCGCAATTCGAACACCGGTTCCAGAAATCGTTTGGAAAGAGTTGG	720
Db	684	GAAAGCGCATTCCTGAGTTGGCACTTAAGATTCAGAGATGTGTTTGGAGAGAGTTGG	743
Qy	721	GGTGACACCGCAGAGCGGTGCTCGAGATGATCAACTCTTTTGGATCTTCTTGAGGCA	780
Db	744	GGTATATCTGTGAGCGCTGCACTTGAGATGATTCACCTTCTTGTGATCTCTTAGGCA	803
Qy	781	ACGTATCCTTGACACCTTGAGAAAGTTCTTGAGAAATCCCAAGGTGTTCAATGTGG	840
Db	804	CTGTATCCGTGCACTCTTGAGACTTTCCTTGCGAGAAATCCCAAGGTCTTTAATGTGG	863
Qy	841	AATTCATCTCCCAACGAGTACTTCGCTCAAGACATGTTTGGGGTATCCCGACACGGT	900
Db	864	AATCTTACTCCCAACGGTACTTTCGCTCAAGACATGTTTGGGGTACTCGTATCTGTT	923
Qy	901	GGGCAAGGTTGTTACATTTGAGATCAAGTCCGAGCTTTGGAGAAATGAAATGCTCTCGT	960
Db	924	GGTAGAGTGTGTTACATTTGAGATCAAGTTCGAGCTTGAGAGATGAAATGCTTCTGTT	983
Qy	961	ATTAAGCAACAGAGACTCAACATCAACCCCTCGAATCTCATTTATTACTAGACTTCTCT	1020
Db	984	ATCAAGCAGAGAGACTTGATATCAACCCCTCAATCTCATTTATTAACCGTCTCTCCCA	1043
Qy	1021	GATGCTGTGGAACAACATGCGGTCAACGACTTGAGAAAGTATAGGAAACAGAGCACTGG	1080
Db	1044	GATCAGATGAGCAACATGTGTCAAGCGACTTGAGAAAGTATATGAACTAATGATTTCA	1103
Qy	1081	GATATTCGTGAGTACCTTCAGAAAGAAAGGAAATTTGTCGAAAAATGATCTCAAGA	1140
Db	1104	GATATTCGTGAGTACCTTCTTTAGACTGAGAAAGGAGATGTCCGACAGTGAATTCACGA	1163
Qy	1141	TTTGAAGAAAGTGTGGCATATCTGAAACCTTACAGAGAGATGTGCTCATGAAATCTCC	1200
Db	1164	TTTGT---AAGTGTGGCTTATCTTAGAGACTTACACCGAGAGATGTGCTGTGAAATTCGC	1220
Qy	1201	AAAGAGTTTGCACGGCAACCGCAGATCTGATCTGAAACAAACAGCAGACGGCAATATGCTC	1260
Db	1221	AAAGAGTTTGCACGGCAACCGTGAATCTAATCATATTTGAAACTTACAGATGATGGCAACTTGT	1280

Qy	1261	GCTCTCTGCTGCACTAATAATAGTGTCAACAGTGCACATCGCCCAATGCTTTGGAG	1320
Db	1281	GCTTCTGTGTAGCACTAATCTAGGTGTACCAAGTGTACTAATGTGCATGCTTTGGAG	1340
Qy	1321	AAGACAAATATTCAGATTCAGATATCTAATTTGAGAAAGCTTGAAGCAATATCAATTC	1380
Db	1341	AAACAAATAATTCAGATTCAGATATCTAATTTGAGAAATTTGAGATGACAAATACACTTC	1400
Qy	1381	TCCTGCAATTTACAGCTGATCTTTTTCAGTAAGAACATACAGATTTTCATCAACAGT	1440
Db	1401	TCATGCAATTTTACAGCTGATCTTTTACAGTGAATGCACTGACTTTCATCAACAGT	1460
Qy	1441	ACTTTCAGAGAAATTTGAGAGAGAGAGACATGTTGTCAATACAGAGCCACACTGCT	1500
Db	1461	ACCTTCAGAGAAATTTGAGAGAGAGAGACATGTTGTGATGAGAGTCAACACTGCT	1520
Qy	1501	TTCACTCTTCCTGTGTCTACCGGTGTGTAAATGTAATGATATGATTTGATATCCCAATTC	1560
Db	1521	TTCACTCTTCCTGTGTCTACCGGTGTGTAAATGTAATGATATGATTTGATATCCCAATTC	1580
Qy	1561	AACATGTTTCCCTGTGTGTGTATGAGATATCTTCCCTTACACCGAAGAGAACGG	1620
Db	1581	AACATGTTATCTCTGTGTGTGTATGAGATATCTTCCCTTACACCGAAGAGAACGA	1640
Qy	1621	AGGTGAGCAATTTTCATCTGAGATGCAAGACCTTCTTAACCCAAAGTTGAGATGAA	1680
Db	1641	AGGTGAGCAATTTTCATCTGAGATGCAAGACCTTCTTAACCCAAAGTTGAGATGAA	1700
Qy	1681	GAAACCTAATGTGTGTCAATGACCGGCAACAGCAATTCGTTCAACAATGCCAAGCTT	1740
Db	1701	GAAACCTAATGTGTGTCAATGACCGGCAACAGCCACTCTGTTTCAACAATGCCAAGTGTG	1760
Qy	1741	GATCGTGTCAAGAACTTAAACCGGACTCGTCAAGTGTGTGCGGCAAGAACCCAAATGTGCT	1800
Db	1761	GACCGTGTAAAGAAATTTGACTGTGTGTGTGAGATGTGTATGGAAGAAATGCTAACGT	1820
Qy	1801	GAGTTGGCTAACCTGT	1860
Db	1821	GAGTTGTGAACCTGT	1880
Qy	1861	GAGAGGCTGAATATGAAGAAATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1920
Db	1881	GAGAGGCTGAATATGAAGAAATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1940
Qy	1921	TTCAATGTGTATCATCTCAATGAAACAAATCCGAAATGTTGAATTTTACCGATCAAT	1980
Db	1941	TTCCGATGTGTCTTCTCAGATGAATGGGCTCAGAAACGGTGAATTTTACCGGATCAAT	2000
Qy	1981	TGGCAGAGAAAGGTGCTTTGTACAGCTGTGATGTGTATGAAAGCTTTGATGATGACGTT	2040
Db	2001	TGTGAACAAAGGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2060
Qy	2041	GTTGAGGCAATGACTTGTGGCTTTTGTCAACATTCGCAACTGTGAAACGGTGTGACAGCGAG	2100
Db	2061	GTTGAGGCAATGACTTGTGATGTGCAACATTTGTCAACTGTGAAAGGAGGCGCGCTGAG	2120
Qy	2101	ATTATGTGTCAATGAGAAATCTGTGTTCAA	2160
Db	2121	ATTATGTGTAAAGGAAATCTGTGTTCAACATTTATCTTAACATGTGTGAGCAGGCTGT	2180
Qy	2161	GACATACTGTGATTTCTTTGAAAGTGTAAAGAAAGATCCATCTCACTGGGATTAATGATC	2220
Db	2181	GAGATCTTGTGATCTTCTTTGAGAAAGTGCAGAGCAAGATCCCTCATATCTGGGACAAATATA	2240
Qy	2221	TCCCAAGAGGCTTGAACGAATAGAGAGAGTATACATGAAGATTTACTGGAAGAGA	2280
Db	2241	TCATGTGGGTGGCCTGAAGCGTATGTAAGAGAGATATCATATGAAATTTTACTCAGAGG	2300
Qy	2281	CTATTGAACCTGTACAGAGGTGTATGGAATTCGGAAGCATGTTTTCACACTTTGAACGGCGT	2340
Db	2301	CTTGTGACCTTGTACGTGAGGTCTATGAGTTCTGTGAACCAATGTATTAACCTTGAATGGCTTT	2360
Qy	2341	GAGAGTGTGTACTTGAAGATGTTTATGTCTTTATGTAAGCTGTAACTGTGGCTGAATCA	2400


```

|||||
Db 2361 GAGAGCCCTCCATACCTGGAATGTTTATGCTCAAGTACCGTAAGCTGGCTGAATCT 2420
Oy 2401 GTTCATGGGAGAGAGTAAATGAACCTGTTAAATGAATGAGCCGGTTCCTGG 2460
Db 2421 GTTCCTGGGCTGGAGTAAATGACATGAGAAAGCTTCAGGGTGGTGGTCCGAGT 2480
Oy 2461 AGAATTAATATCTGTTTGTGA 2481
Db 2481 GGAATTTCTCGAGTTGTGGA 2501

RESULT 3
VIRVSS1
LOCUS 2652 bp mRNA linear PLN 02-FEB-1999
DEFINITION Vigna radiata mRNA for sucrose synthase, complete cds.
ACCESSION D10266
VERSION D10266.1 GI:218332
KEYWORDS D-fructose 2-glucosyltransferase; UDPglucose; sucrose synthase;
vsa1.
SOURCE Vigna radiata
ORGANISM Vigna radiata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Vigna.
REFERENCE 1 (bases 1 to 2652)
AUTHORS Arai,M., Mori,H. and Imaseki,H.
TITLE Expression of the gene for sucrose synthase during growth of mung
bean seedlings
JOURNAL Plant Cell Physiol. 33, 503-506 (1992)
COMMENT Submitted (21-Jan-1992) to DBJ by:
Masao Arai
National Institute for Basic Biology
Myodaiji-cho, Okazaki
Japan
Phone: 0564-55-7611
Fax: 0564-53-7400.
FEATURES
Source Location/Qualifiers
1..2652
/organism="Vigna radiata"
/mol_type="mRNA"
/db_xref="taxon:157791"
/cclone="Vsa1"
/cisue_type="stem"
/dev_stage="seedling"
1..2652
/gene="vsa1"
38..2455
/gene="vsa1"
/EC_number="2.4.1.13"
/codon_start=1
/product="sucrose synthase"
/protein_id="BA01108.1"
/db_xref="GI:218333"
/translacion="MARDRLRVSHSRDLDTLSANRETLALSLRBEKSKILQH
HOVAEFBEIPESRKOITDGAFGSVLSTDEALVLPVVALARPFGWEYLRVW
HALVVEVLPAPAEVLRFEKELVDSSNGFVLELDEPPTASFPPTLANKSINGVOFL
NRHLNAXLFDHKEISLPLELRLSHVSGKTLMDRLIONDALQHVLRKAEYLGTV
PPTPYSAFBEHFOEIGLERGMNAEVLBSIOLDLLEAPDCTLETFGLPMV
FNVLISPHGVFADNVLYGPDGGOVYLLDOVALDENMLHRIKOGLDIVRIILI
ITRLPRAVCTTCQRLEKVRGTEHSHILRVPPRTENGIVAKWISREWPYLETYS
DVAMELAKEDQKFDLIVGNSDGNIVASLAKLQVTOCTIAHALEKTKIPSDIYV
KLEERHFSQPTADLPANMTDPIITSTQELAGSDPTVGQYESHTAFTLPGLYV
VGEIDVDPKFNIVSPGADQTIYPHTEISRLLTSFHELELLSYSENEHICVVK
DRSKPIITMARLDRKVNITGLVEMGNALRELNVLVVAGDRRKESKDLERKAM
KQWYSLIETVTLNGOFWISOMRVNGELRYIADTKGAFVQPAYEARGLIVVA
MTGCLPTPATCNGSPAEIIVYKSGFHDIPHGRRADLIVPEFKYKVDPSHDKXS
OAGLRIEKTMOIYSORLLITLGVGFVHVSNDLRRESRYLEMFALTKYKLA
SVPLAVE"
2652
/gene="vsa1"
polya_site

```

```

ORIGIN
Query Match 61.5%; Score 1615.6; DB 8; Length 2652;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 1931; Conservative 0; Mismatches 505; Indels 3; Gaps 1;

Oy 1 ATGGCTGAGCGTGTCTCATCTCGGCTCCAGATCTCCGTGAGCGTTTGATGAGACCTT 60
Db 38 ATGGCTACCGATCGTTTGACCCCGTGTTCAGAGTCTCCGTGAGAGCTTGATGAACCTTG 97
Oy 61 CTGTGCTACAGAAACGAGATTTTG3CCTTGCTCTCAGAGATCGAGGCAAGAAAAGGA 120
Db 98 TCTGCAACAGAAAGAAATTTCTG3CCTTGCTCTCAGAGATCGAAGCAAGGCAAGGGG 157
Oy 121 ATTCTGCAACACCTCAATTAATTTAGAGTTTGAAGTATTCCTGAAAGAAACAGAAAG 180
Db 158 ATTTTGCAACACATCAGATCATTTGCGAGATTTGAGAAATCCCGAGAGACAGACAG 217
Oy 181 AAGCTCGCTAATGATGCAATTTTGAAGTATTTGAAGCTAGTCAGAAAGCATGCTGTG 240
Db 218 AAGCTTACTGATGATGCTTGTGAGAAATTTGAGATCTACTCAGAAAGCATGATTTTG 277
Oy 241 CTTCCATGAGTGTGACATTTGCTGTTCAGAGCGCTGTGTGAGTACATTAGAGTG 300
Db 278 CCACCATGAGTGTGCTGTGCGCTGTTCAGAGCGCTGTGTGAGTACATTAGAGTG 337
Oy 301 AATGTCACGCGCTTGTGTGTGAGAACTCATGCTGTGTGATCTCCACTTCAAGGA 360
Db 338 AATGTGACGCTGTGATGTGTGAGAGTGTGCACTGTGTGATCTCCGCTTCAAGAG 397
Oy 361 GAGCTTGTGATGAGAGTCAATGAATGAATCTTGTGTGAATGATTTGAGCGCTTC 420
Db 398 GAATCTGTGATGAGAGTCTAATGAGCACTTGTGTGAGTGTGAGTGTGAGTGTGAG 457
Oy 421 AACTCATATTCGCCCGCCCACTCTTCAAAATCCATGTGTAAATGTGTGAGTGTCTA 480
Db 458 ACCGATCTCTTCCCGCCCACTCTTCAAAATCCATGTGTAAATGTGTGAGTGTCTC 517
Oy 481 AATGTCACCTTGTGTGAGAAATTTGTCATGACAGAGAGATGCAACCTTGTGTGGA 540
Db 518 AATGTCACCTTGTGTGAGAAATTTGTCATGACAGAGAGATGCAACCTTGTGTGGA 577
Oy 541 TTCTCATGAGTCAATGTGCAAGAGGCAAGATGATGATGATGATGATGATGATGATG 600
Db 578 TTCTCATGAGTCAAGAGTCAAGAGGCAAGATGATGATGATGATGATGATGATGATG 637
Oy 601 TTGAATGCTCTTCAACATGTTTGTGAGAAAGAGAGATGATGATGATGATGATGATG 660
Db 638 CCGATGCTCTTCAACATGTTTGTGAGAAAGAGAGATGATGATGATGATGATGATGATG 697
Oy 661 GAGACACCATGTGCGGAATTTGAAACACCGGTTCCAGAAATCGGTTTGAAGAGGTTGG 720
Db 698 GAAACCCCTTACTAGATTTGAGACACAGATTCAGAGATTTGAGAGAGGTTGG 757
Oy 721 GGTGACACCGGAGAAAGGCTGAGATGATGATGATGATGATGATGATGATGATGATG 780
Db 758 GGTGACACCGGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 817
Oy 781 ACTGATCTTGTGACCTTGTGAGAAATTTGTTGAGAAATCCCATGATGATGATGATG 840
Db 818 CCGATCTTGTGACCTTGTGAGAAATTTGTTGAGAAATCCCATGATGATGATGATGATG 877
Oy 841 ATTCTCACTCCCAAGGATTTGCTGTCAGAGCAATGTTTGGGGTATCCGCAACCGGT 900
Db 878 ATTCTTCTCCCAAGGATTTGCTGTCAGAGCAATGTTTGGGGTATCCGCAACCGGT 937
Oy 901 GGGGAGGTGTTTACATCTTGATCAAGTCGAGCTTGTGAGAAATGATGATGATGATGATG 960
Db 938 GGTAGGTTGTTTACATCTTGATCAAGTCGAGCTTGTGAGAAATGATGATGATGATGATG 997
Oy 961 ATTAAGCAACAGAGACTCAACATCACCTCGATCTCAATTAATTAAGCACTTCTTCT 1020
Db 998 ATTAAGCAACAGAGACTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1057

```

QY 1021 GATCTGTGGAACAATGCGGTGCAAGCACTTGAGAAAGTATACGGAACGAGCACTCG 1080
 DB 1058 GATCAGAGTAGTACTTTCGGCCCAAGCTTGAGAAAGTGTGGAACTGAACTACTCT 1117
 QY 1081 GATATTCCTGAGAACCTTTCGAAACGAAAGGAATGTTCCAAAATGGAATCTCAAGA 1140
 DB 1118 CACATTTCTACAGATTCCTTTAGAACTGAGAACGAAATGTTCCAGATGATCTCAAGA 1177
 QY 1141 TTTGAAAAAGCTGAGCCATTACTTGAAACCTACACAGAGATGTTGCTCATGAATCTCC 1200
 DB 1178 TTCC---AAGTCTGGCCCTTACTTGAAACCTACACAGAGATGTTGCTCATGAATCTGCA 1234
 QY 1201 AAAGATTGACGCGACGCCAGATCTGATTCGAAACGAAACGACGAGCAATATGCTC 1260
 DB 1235 AAAGATTGACGCGACGCCAGATCTGATTCGAAACCTACAGTATGAAACATTTGT 1294
 QY 1261 GCCCTCTTGTGCGACATTAATTTAGTGTCAACAGTGCACATTCGCCCATCTTTGGAG 1320
 DB 1295 GCCCTTGTGTGCGACATTAATTTAGTGTCACTAGTGTACATTTGCTCATGCACTTGAG 1354
 QY 1321 AAGACAAATATCCAGATTTCAGATATCTATTGAAAGAGCTTGAAAGCAATATCATTTTC 1380
 DB 1355 AAGACAAATATCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1414
 QY 1381 TCTTGCCAAATTTCAGCTGATCTTTTTCGAAATGAACATATGATTCATCATCAGCAGT 1440
 DB 1415 TCTTGCCAAATTTCAGCTGATCTTTTTCGAAATGAACATATGATTCATCATCAGCAGT 1474
 QY 1441 ACTTTCCAGAAATTTGCGAAGAGCAAGCACTGTTGCTCAATACGAGAGCAACTGCT 1500
 DB 1475 ACATTTCCAGAAATTTGCGAAGAGCAAGCACTGTTGCTCAATACGAGAGCAACTGCT 1534
 QY 1501 TTCACTCTCTGCTGATCTTACCGTGTGTATCATGATTCATGATTCATTCCTCAAAATTC 1560
 DB 1535 TTCACTCTCTGCTGATCTTACCGTGTGTATCATGATTCATGATTCATTCCTCAAAATTC 1594
 QY 1561 AACATTTGTTCCCTGCTGATTCAGATATGAGATATCTTCTCTTACCCGGAAGAGAGCGG 1620
 DB 1595 AACATTTGCTCTCCCGAGCTGATTCAAACATTTACTTCCCCACACGGAATATGCGCG 1654
 QY 1621 AGCTTGAAGCAATTTCCATCTGATTCAGATTCAGATTCATTCATTCATTCATTCATTCAT 1680
 DB 1655 AGCTTGAAGCAATTTCCATCTGATTCAGATTCAGATTCATTCATTCATTCATTCATTCAT 1714
 QY 1681 GAACACTATGCTGCTGATTCAGATTCAGATTCAGATTCATTCATTCATTCATTCATTCATTCAT 1740
 DB 1715 GAACACTATGCTGCTGATTCAGATTCAGATTCAGATTCATTCATTCATTCATTCATTCATTCAT 1774
 QY 1741 GATGCTGTCAAGAACTTAAACGGAATCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1800
 DB 1775 GATGCTGTCAAGAACTTAAACGGAATCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1834
 QY 1801 GATGCTGTCAAGAACTTAAACGGAATCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1860
 DB 1835 GATGCTGTCAAGAACTTAAACGGAATCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1894
 QY 1861 GAGAGAGCTGAATGAAGAAATGTTGAGCTGATTCGACAGTAACTTGAACGCGCAA 1920
 DB 1895 GAGAGAGCTGAATGAAGAAATGTTGAGCTGATTCGACAGTAACTTGAACGCGCAA 1954
 QY 1921 TTTGAGATGATATCATCTCAATGAACGAAATCCGAAATGTTGAACCTTACCGATACAT 1980
 DB 1955 TTTGAGATGATATCATCTCAATGAACGAAATCCGAAATGTTGAACCTTACCGATACAT 2014
 QY 1981 TTTGAGATGATATCATCTCAATGAACGAAATCCGAAATGTTGAACCTTACCGATACAT 2040
 DB 2015 TTTGAGATGATATCATCTCAATGAACGAAATCCGAAATGTTGAACCTTACCGATACAT 2074
 QY 2041 GTGAGAGCAATGATCTGCTGTTGCGCAATTCGCAACTTGAACGCTGAGCAAGCGAG 2100
 DB 2075 GTGAGAGCAATGATCTGCTGTTGCGCAATTCGCAACTTGAACGCTGAGCAAGCGAG 2134

QY 2101 ATTAATGTCATGAGAAATTCGTTTCAACATTGATCTTCAACATGCTGATCAAGCTGCT 2160
 DB 2135 ATATGCTGATGAGAAATTCGTTTCAACATTGATCTTCAACATGCTGATCAAGCTGCT 2194
 QY 2161 GACATCTGCTGATTCCTTTGAAAGTGTAAAGATGATTCATCTGATGATGATGATGATGATGAT 2220
 DB 2195 GACCTCTGTTGATTCCTTTGAAAGTGTAAAGATGATTCATCTGATGATGATGATGATGATGAT 2254
 QY 2221 TCCGAGAGCTTGAAGCAATGAGAGATGATTCATGAGATGATTCATCTGAGAGAGA 2280
 DB 2255 TCCGAGAGCTTGAAGCAATGAGAGATGATTCATGAGATGATTCATCTGAGAGAGA 2314
 QY 2281 CTATGACCTGACAGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2340
 DB 2315 CTATGACCTGACAGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2374
 QY 2341 GAGAGCTGCTGATTCATGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2400
 DB 2375 GAGAGCTGCTGATTCATGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2434
 QY 2401 GTTCCATTGCGAGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2439
 DB 2435 GTTCCATTGCGAGAGATGATTCATGAGATGATTCATGAGATGATTCATGAGAGAGA 2473

RESULT 4
 AF030231
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 2842)
 Zhang,X.-Q., Verma,D.P.S., Patil,S., Arredondo-Peter,R.,
 Miao,G.-H., Kuismanen,R., Klucas,R.V. and Chollet,R.
 Cloning of a full-length sucrose synthase cDNA from soybean
 (Glycine max) root nodules (Accession No. AF030231) (PGR 97-173)
 Plant Physiol. 115, 1729 (1997)
 2 (bases 1 to 2842)
 Zhang,X.-Q., Verma,D.P.S., Patil,S., Arredondo-Peter,R.,
 Miao,G.-H., Kuismanen,R., Klucas,R. and Chollet,R.
 Direct Submission
 Submitted (16-OCT-1997) Biochemistry, University of
 Nebraska-Lincoln, George W. Beadle Center, Lincoln, NE 68588, USA
 Location/Qualifiers
 1..2842
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /feature_type="nucleotide" root nodules"
 1..2842
 /gene="SS"
 /BC_number="2.4.1.13"
 /note="nucleotide-100"
 /codon_start=1
 /product="sucrose synthase"
 /protein_id="AAC39323.1"
 /db_xref="GI:2606081"
 /translation="MATHDRTVHSRBERLDELTLNRENIALLSRLEKAGKGILOH
 HQVIAFEBELPEENRQKLTIDGAFGEVLRSTQEAIVLPWVALVPRPGWEYLRVAV
 HALVVELQPAEYLFKEBELVDSSNGNFTLEIDFEPNPAFPRPTLNKISGNQVLT
 NRIISAFLFDKESLFLLEFLRLHSVKGKTLNDRIONPDALQVLRRAEYLGTV
 PEPPTVSEPERFOELGLERGWGDNARVLESLQILDLEADPCTLETFLEGIPIV
 FNVIISPHGYRQDNVLTGVRPDGQGVVYLLDOVRLALENMRIRKIOGGIDYRLLI
 ITRLPDVAGTTCGRLEKVFTEHSHIILAVPRTEKGIYRKNIISREVPYLETYTE

DNAAELAKELQKBDLIVGNYSIDGNIVASLHAKLVTOCTIAHALEKTKYBSBDIYM
KLEERHPSGCTPADLFAAMHDTPIITSTFOELIAGSDOTGQESHTAFLPELJYV
VHGIDVDPKKNIVSPGADORTIYPHPETISRLSPHEBELLSYSENEHI CVLK
DRSRIPIIETMARLRKRNITGLVEMYGNAULRELVNIVYVAGSRRESNDLERAKM
KMTGLIETIKLNGQFRNISQMKRVRNDELRYVICTRGNFVPAYAEALGTVEA
MTGLPTATCNGBPAEIVHGKGFHIDPHGRADLDVDFEKKDLPTMDKXS
KAGIORIBEKTYMOIVISORLLITLNGVGFHMGVSNLDRBSRYLMEFYALKYKRLAR
SVPLAAB"

ORIGIN

Query Match 61.1%; Score 1602.6; DB 8; Length 2842;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 1918; Conservative 0; Mismatches 505; Indels 3; Gaps 1;

1 ATGGCTGAGCGGTCTCACTCGGCTGCAAGTCTCCGTGAGCGCTTGATGAGCCCTT 60
Db ATGGCCACCGATCGTTTGAACCCGGGTTCACAGTCTCCGTGAGAGCTTGAGAAACCTCTC 131
Qy CTTCCTCACAAGAAAGAGATTTTGGCTTGCCTCAGAGATCGAGGCGAAGAAAGAA 120
Db ACTGCGAACAAGAAATTTTGGCCCTTCTGTCAAGATCGAAGCCAGGCGAAGGCG 191
Qy 121 ATTCTGCAACCAATCAATATTATCTAGAGTTTGAAGCTATCCCTGAAGAAACAGAAAG 180
Db 192 ATCTGCAACCAACCAAGGTCAATGCTGATTTGAGAAATCCCTGAGAGAAACAGACAG 251
Qy 181 AAGCTCGTATGCTGCAATTTTGAAGATTTGAAGCTAGTCAGAAAGCAATGCTTGT 240
Db 252 AAGCTCACTGATGCTGCTTGGAGAAAGCTTGAGATCTACAGAGAAACCAATGTTTG 311
Qy 241 CCTCCATGAGTTGCACTTGTCTGCTCAAGAGCGTGTGTTGGAGATCAATTAAGATG 300
Db 312 CCACCAATGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 371
Qy 301 AATGTTCAAGCGCTTGTGTGAGAGAACTGCTGTGAGATCTCCACTTCAAGAA 360
Db 372 AATGTTCAAGCGCTTGTGTGAGAGATGAGAACTGCTGAGATCTCCACTTCAAGAA 431
Qy 361 GAGCTTGTGATGAGATTTCAATGAGAACTTTGTTTGAATTTGAATTTGAGCCCTTC 420
Db 432 GAACCTGTGAGAGAAAGTTCAATGAGAACTTTGCTGTGAGTTGAGCTTTGAACCAATTC 491
Qy 421 AACTCATCATTTCCCGCCCAACTCTTCAAAATCATTTGTAATGAGTGAAGTCTTA 480
Db 492 AATGAGCGCTTCCCGCCCAACTCTTCAAAATCATTTGTAATGAGTGAAGTCTTA 551
Qy 481 AATGCTCACTTTGCGCAAAATGTTCCATGACAAAGAGATGACACCTTTGCTGCA 540
Db 552 AACGTCACCTTTCTGCAAACTCTTCAAGCAAAAGAGATGACACCTTTGCTGCA 611
Qy 541 TTCTCTCAGATCTCATGTCACAAAGGCAAGAAATGTTGATGACAGAAATTCAGAAC 600
Db 612 TTCTCTCAGATCTCATGTCACAAAGGCAAGAAATGTTGATGACAGAAATTCAGAAC 671
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTGATCCCTACTCTCT 660
Db 672 CCAATGACATCTCAACATGTTTGAAGAAAGAGATCTTGTGATCCCTACTCTCT 731
Qy 661 GAGACACCATGTCGGAATTCGAACACCGGTTCCAGAAATCGTTTGAAGAAAGTTGG 720
Db 732 GAAATCCCTACTGGAATTTGAGCACAAGTTCCAGAGATTTGGTTTGAAGAGAGGTTG 791
Qy 721 GGTGACACCGCAAGACGCTGCTGAGATGATCCAACTCTTTTGAATCTTTTGAAGCA 780
Db 792 GGTGACACCGCAAGACGCTGCTGAGATGATCCAACTCTTTTGAATCTTTTGAAGCC 851
Qy 781 ACTGATCTGTCACACCTTGAAGAAAGTTCTTGGAGAAATCCCAATGATGATGATG 840
Db 852 CTAACCCGTCACACCTTGAAGATTTCTTGGAGAAATCCCAATGATGATGATGATG 911
Qy 841 ATTCTCACTTCCCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 912 ATTCTTCTTCCCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATG 971

Qy 901 GGGCAGGTTGTTTACATCTTGATCAAGTCGAGCTTTGAGATGAGATGCTCTCCGT 960
Db 972 GGGCAGGTTGTTTACATCTTGATCAAGTCGAGCTTTGAGATGAGATGCTCTCCGT 1031
Qy 961 ATAAAGCAACAGGACTCAACATCAACCCCTGAAATCTCATTAATTAATTAATTAATTA 1020
Db 1032 ATTAAGCAACAGGATTTGACATGTTCTCTGATTTCTCATTAATTAATTAATTAATTA 1091
Qy 1021 GATCTGTGGAACAACATGCGGTCAACGATTTGAGAAAGATTAACGGAACGACACCTCG 1080
Db 1092 GATCAGATGGAACCTATCTTGGGCAACGCTTTGAGAAAGTTGTCGAACTGACACCTCC 1151
Qy 1081 GATATCTTGAATGACCTTCAAGAAAGAAAGGAAATTTGTCGAAATGATCTCAAGA 1140
Db 1152 CACATTTCTGAGATTTCTTTTGAATGAGAAAGATTTGTCGAAATGATCTCAAGA 1211
Qy 1141 TTTGAAAAAGCTTGGCAATCTTGAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
Db 1212 TTGG--AAGTCTGCGCTTACTTGAACCTTACAGAGATGTTGCTCATGAATCTCC 1268
Qy 1201 AAAAGTTGCAAGGCAAGCGCATCTGATCATGGAACCAACAGCGCAATATGCTC 1260
Db 1269 AAAAGTTGCAAGGCAAGCGCATCTGATCATGGAACCAACAGCGCAATATGCTC 1328
Qy 1261 GCTCTCTGCTGCAATTAATTTAGGTGTCACAGTGCACCATGCGCCATGCTTGGAG 1320
Db 1329 GCTCTTGTGTGCAATTAATTTAGGTGTCACAGTGCACCATGCTTGGAG 1388
Qy 1321 AAGCAAAATATCCAGATTCAGATATCTATGGAAGAGCTTGAAGCAATATCAATTC 1380
Db 1389 AAGCAAAATATCCAGATTCAGATATCTATGGAAGAGCTTGAAGCAATATCAATTC 1448
Qy 1381 TCTTGCATTTAAGCTGATCTTTTGAATGAAACCAATATTCATATCAACAGT 1440
Db 1449 TCTTGCATTTAAGCTGATCTTTTGAATGAAACCAATATTCATATCAACAGT 1508
Qy 1441 ACTTTCAGAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1509 ACTTTCAG 1568
Qy 1501 TTCACTCTCTGCTCTCTACCGTGTGTCATGATATGATATGATATGATATGATATGAT 1560
Db 1569 TTCACTCTCTGCTCTCTACCGTGTGTCATGATATGATATGATATGATATGATATGAT 1628
Qy 1561 AACATGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAAGAGAGAG 1620
Db 1629 AACATGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAAGAGAGAG 1688
Qy 1621 AGGTGGAAGATTTCCATCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1689 AGGTGGAAGATTTCCATCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1748
Qy 1681 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1749 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
Qy 1741 GATGCTGTCAGAACTTACCGGATCTGTCAGATGTCGCGCAAGAAACCAATGTCGT 1800
Db 1809 GATGCTGTCAGAACTTACCGGATCTGTCAGATGTCGCGCAAGAAACCAATGTCGT 1868
Qy 1801 GATGCTGTCAGAACTTACCGGATCTGTCAGATGTCGCGCAAGAAACCAATGTCGT 1860
Db 1869 GATGCTGTCAGAACTTACCGGATCTGTCAGATGTCGCGCAAGAAACCAATGTCGT 1928
Qy 1861 GAGAGGCTGGAATGAGAAATGTTTGAAGCTGATGACAACTTGAACGCGCA 1920
Db 1929 GAGAGGCTGGAATGAGAAATGTTTGAAGCTGATGACAACTTGAACGCGCA 1988
Qy 1921 TTCAATGATATATCTCAATGACAGAAATCCGAATGTTGAATTTAATCAATTAAT 1980
Db 1989 TTCAATGATATATCTCAATGACAGAAATCCGAATGTTGAATTTAATCAATTAAT 2048

QY	1981	TGCACACAAAGGTCCTTGTACAGCTGCATTGATGAAGCCTTGGATTGACAGTT	2040
Db	2049	TGCACACACAGGGGTCTTTCGTGTGACGCTCTGTATACAGGCTTTTGTGTGACAGTG	2108
QY	2041	GTGAGGCAGATGACTTGTGGGTTTCCCAACATTGCGAACCTGTAAAGGTGACACAGCCGAG	2100
Db	2109	GTTGAGGCCAGATGACTTGTGGGCTTCCCAACATTCCCAACATGCAATGTGTGTCTCTGTAGG	2168
QY	2101	ATTATTGTCCATGGGAATCTGTGTTTAAACATTATCTTATCCATGTGATTCAGAGTCGT	2160
Db	2169	ATCATGTGACCGGCAAGCTTGGCTTCCACATTACCTTTACAGATGACCGTGTGCT	2228
QY	2161	GACACTCTGTGCAATTTCTTTGAAAGGTATGAAGAAAGATCCATCTCACTGGGATTAAGATC	2220
Db	2229	GATCTCTGTTGACTTCTTTGAGAAGTGCAGGCTTGACCACTCACTGGGACAGATATC	2288
QY	2221	TCCCAAGAGGCTTGGAAACGAATAGAGAGAAAGTATACATGGAAGATTACTCGGAAGAGA	2280
Db	2289	TCAAAGGCTGTCTCCAGCGCTGATTTGAAGAGAAAGTACATGGCAAAATTTACTTCBAGG	2348
QY	2281	CTATTGACCGCTGACAGGAGTGTATGGAATCTTGGAAGCAATGTTTCCAACCTTGAACGCCGT	2340
Db	2349	CTTCTCACTCTCACCGGTGTCTATGTGCTTCTGGAAGCAGATGTGTAACTTGAACGCCGT	2408
QY	2341	GAGAGTCGTGTTAACCCTTGAGATGTTTTATGCTCTTAAATACCGTAAAGCTGGCTGAATCA	2400
Db	2409	GAGAGCGCGGCTATCTCGAGATGTTCTATGCTCTCAAGTACCGCAAAATGGCTGAAGTCT	2468
QY	2401	GTTTCATTTGGACAGAGAGTAATTGA	2426
Db	2469	GTGGCCCTTGTGCTGAATTAACCTGA	2494

RESULT 5	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
AF315375	AF315375	Phaseolus vulgaris sucrose synthase mRNA, complete cds.	2738 bp	GI:25990359	Phaseolus vulgaris	Phaseolus vulgaris	1 (bases 1 to 2738)	Camas,A., Silvente,S. and Lara,M.	Expression of the gene for sucrose synthase during development of common bean nodules	Unpublished	2 (bases 1 to 2738)	Camas,A., Silvente,S. and Lara,M.	Direct Submission	Submitted (20-OCT-2000) Programa de Biologia Molecular de Plantas,

FEATURES	
source	Universidad S/N. Cuernavaca, Mor. 62210, Mexico
	Location/Qualifiers
	1. .2738

CDS

```

/organism="Phaseolus vulgaris"
/mol_type="mRNA"
/db_xref="taxon:3885"
/clone="PVSS"
/class_type="modules"
52..2469
/EC_number="2.4.1.13"
/notes="SS; D-fructose 2-glucosyltransferase; UDP glucose;
sucrose cleavage; predicted 92 kDa"
/codon_start=1
/product="sucrose synthase"
/protein_id="AA06498.1"
/db_xref="GI:25990360"
/translation="MAADRTTRVHSIRELDSTLSANRRELIALISLEAKGICILQ

```

HUIVAEFEEBEPESRQKLDIDGAGEVLRSQOELIAPPMWALVPRPPGMEWLVAV
 HALVAVELDPAEYLAKREBLVDSGNSFNFLDEPFBTPASPRITNKSINGVOLF
 NHRLSAKLFRIDKESLHPLFLRLHSYNGKTLMLNDRIQNPDLQHYVRAQEEYLGIV
 PRETPYSEFBHKFQDIGLERGMGANBLVLESIQLLDLLEADDPCTULEFPLIRLIPW
 FNVVLSPHGYFAQDNVLGYPDPTGGQVVYLIDOVRLLENMLHARKOQGLIVPRILI
 ITLLELDVAGTTCGRLKEKVFGEHSHILAVRPRTENGRIKALISEFVWPLYETYE
 DVHLELAKELQGGKPDILGVNGSDNIVASLHKLQVTOCTILHAKELTSPSDIYV
 KKLDEERYKSCOPTADLFAMNHDPDITISFPOSLASKDVIQVGEYSHATPLAGLIVAV
 VAGIDVDFPREFTVSPGADOTIYFSPKESRITLTHSPREILVLESVENHEBICVAK
 DRKPDILFTMARLDVRKNIPLGVLEWYGKQEKRELVNLVAVAADRRRESDELEKEM
 KKNVSLIETVYKLGNGCFRMWISSOMNRRVNGELVYVIDTRCAFOVAPAYAFGLTVEA
 MTCGLFPATCNGSPAEILIVHGKSGFHHIPEYHDRAADLIVFEFCKVSPSHMDITS
 QAGLORIEKEKYTWQIYSQRLTLTIVGYGFKHVSJNDRLESRRYLEKVFALXKTKLAE
 SYVPLANE"

Query Match	60.1%	Score 1577.2	DB 8	Length 2738
Best Local Similarity	78.3%	Pred. No. 0		
Matches 1904	Conservative 0	Mismatches 524	Indels 3	Gaps 1
QY	1	ATGGCTGAGGCGTGTCTCACTGCGCTCCACAGTCTCCGTGAGCGTTTGGATGAGACCTTT	60	
DB	52	ATGGCGCGTGATGTTTGAGCCCGTGTTCACAGTCTCCGTGAGGCGCTTGATGAAACCTTG	111	
QY	61	CTTGCTCACAGAAACGAGATTTTGGCCCTTGCTCTCAAGATTCAGAGGCAAGAAAAGGA	120	
DB	112	TCGTCCACAGAAACGAAATTTCTGGCCCTTCTGTCAAGATTAAGCCAAAGGCGACGGGG	171	
QY	121	ATTCTGCACACCATCAATTAATTTAGATTGAAAGTATCCCTGAAGAAACAGAAAG	180	
DB	172	ATCTTGCACACCATCAGGTCAATTTGCTGAGTTGAGAAATTTCCCGAGAGACACAGAC	231	
QY	181	AAGCTGCTTAATGATGATTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGA	240	
DB	232	AAGCTTAATGATGATGCTTGTGAGAAATTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGA	291	
QY	241	CCTTCACAGGTTTCACTGCTGTGTCTCCAGAGCGCTGTGTGTGTGTGTGTGTGTGTGT	300	
DB	292	CCACGATGGTTCTCTGT	351	
QY	301	AATGTTACGCGCTTGT	360	
DB	352	AATGTGACCGCTTGT	411	
QY	361	GAGCTTGTGATGGAAGTTCAATGGAACCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT	420	
DB	412	GAACTGTGTGATGGAAGTTCTTAATGGAACCTTTTGTGTGTGTGTGTGTGTGTGTGTGT	471	
QY	421	AAGCTCATTTCCCGCGCCCACTTTTCAAAATTCATTTGTAATGTTGTGAGTTCTTA	480	
DB	472	ACGGCATTCCTTCCCGCGCCCACTTTTCAAAATTCATTTGTAATGTTGTGAGTTCTTC	531	
QY	481	AATGTCACCTTTTGGCGAAATTTGTTTCATGACAGAGAGCATGACCTTTTGTCTGAA	540	
DB	532	AACCGGCACCTTTTGTCTCAAACTTTTTCATGACAGAGAGCATGACCTTTTGTGAA	591	
QY	541	TTTCTCAGAGTCCATTGTGCAAGAGGCAAGAACATGATGTTGTAATGACAGAAATTCAGAC	600	
DB	592	TTTCTCAGAGTCCATTGTGCAAGAGGCAAGAACATGATGTTGTAATGACAGAAATTCAGAA	651	
QY	601	TTGATGCTCTTCAACATGTTTGAAGAAACGAGAGATTTTGGTATCCTTACTCTCT	660	
DB	652	CTGATGCTCTTCAACATGTTTGAAGAAACGAGAGATTTTGGTATCCTTACTCTCTCT	711	
QY	661	GAGACACCATGTCGCAATTTGCAACACCGGTTCAAGAAATGCGTTTGAAGAAAGGTTGG	720	
DB	712	GAAACCCCTCACTCAAGAAATTTGCAACACCGGTTCAAGAAATGCGTTTGAAGAAAGGTTGG	771	
QY	721	GGTGAACCGCAGAGAGCGGTGCTGAGATGATCAATCTCTTTTGGATCTTTTGAAGCA	780	
DB	772	GGTGAACCGCAGAGAGCGGTGATCTTGAAGTCAATCTCTTTTGGATCTTTTGAAGCC	831	
QY	781	ACTGATCTTTCACCTCTGGAAGTTCTTTGGAGAAATCCCATGCTGTTCAATGTTTGTG	840	

Db 832 CCTATCGTGCACCCCTTGAGACCTTCTTGAGAAATCCCTATGCTTTAAAGTGT 891
 Qy 841 ATTCTCACTCCCAACGATTAATGCTCAACAACAATGTTGGGGTATCCCGAACCGGT 900
 Db 892 ATTTCTTCGCCCCAGTGTACTTTCGCCAAGATATATGCTGGGATACCTCGAACCTGGT 951
 Qy 901 GGGCAGGTTGTTTATCATCTTGATCAAGTCCGAGCTTGGAGAAATGAGATCTCTCCGT 960
 Db 952 GGTACAGTTGTTTATCATCTTGATCAAGTCCGAGCTTGGAGAAATGAGATCTCTCCGT 1011
 Qy 961 ATAAAGCAACAAGACTCAACATCAACCCCTGAAATCTCATTTACTTAACTTCTTCT 1020
 Db 1012 ATTAAGCAACAAGACTTGAATGATGTTCTCTGTAATCTTATTAACAAGCTTCTCCCA 1071
 Qy 1021 GATGCTGCGAACAACATGCGGTCAAGATCTTGAGAAAGTATACGAGAACAGACTCG 1080
 Db 1072 GATCAGAGTGTACTTCTTGAGCAAGTCTTGAGAAAGTGTGGAACCTGAACTGCTCC 1131
 Qy 1081 GATATCTTGAGTACCTTCAAGACAGAAAGGGAATGTTGCAAAATGATCTCAAGA 1140
 Db 1132 CACATCTTGAGTACCTTCAAGACAGAAAGGGAATGTTGCAAAATGATCTCAAGA 1191
 Qy 1141 TTTGAAAAGTCTGCGCATTAATCTTGAAAACCTACACAGAGATGTTGCTCATGAATCTCC 1200
 Db 1192 TTGG---AAGTCTGGCCCTTACTTGGAACCTTACACAGAGATGTTGCTCATGAAGTCC 1248
 Qy 1201 AAAGATTTGACCGGACCGGAGATCTGATCATCGGAAACACAGGACGCAATATGCTC 1260
 Db 1249 AAAGATTTGACCGGACCGGAGATCTGATCATCGGAAACACAGGATGGAACATGTT 1308
 Qy 1261 GGCCTCTTGCGCACTTAATTTAGTGTCAACAGTGCACATCGCCCATGCTTGGAG 1320
 Db 1309 GGCCTCTTGCGCACTTAATTTAGTGTCAACAGTGCACATCGCCCATGCTTGGAG 1368
 Qy 1321 AAGCAAAATATTCAGATTAATCTTATTTGAGAAAGCTTGAAGCAAAATACCATTTTC 1380
 Db 1369 AAGCAAAATATTCAGATTAATCTTATTTGAGAAAGCTTGAAGCAAAATACCATTTTC 1428
 Qy 1381 TCTTGCCAAATTTACAGTGTCTTTTGGAAATGAAACATACATTTATCATCAAGT 1440
 Db 1429 TCTTGCCAAATTTACAGTGTCTTTTGGAAATGAAACATACATTTATCATCAAGT 1488
 Qy 1441 ACTTTCAGAAATTTGAGAAAGCACTGTTGTCATTCGAGGACCACTGCT 1500
 Db 1489 ACATTTCAAGAAATTTGAGAAAGCACTGTTGTCATTCGAGGACCACTGCT 1548
 Qy 1501 TTCACTCTTCTGCTCTTACCGTGTGTTGATCATGATGATGATGATGATGATGATGATG 1560
 Db 1549 TTCACTCTTCTGCTCTTACCGTGTGTTGATCATGATGATGATGATGATGATGATGATG 1608
 Qy 1561 AACATTTCTTCTGCTCTTACCGTGTGTTGATCATGATGATGATGATGATGATGATGATG 1620
 Db 1609 AACATTTCTTCTGCTCTTACCGTGTGTTGATCATGATGATGATGATGATGATGATGATG 1668
 Qy 1621 AGTTTGAAGCATTTTCACTCTGATGATGAAAGCTTCTTACCAACAAAGTTGAAGTAA 1680
 Db 1669 AGTTTGAAGCATTTTCACTCTGATGATGAAAGCTTCTTACCAACAAAGTTGAAGTAA 1728
 Qy 1681 GAACACTATGTTGCTCAATGACCGCAACAAGCAATTTCTTCAACATGCAAGGCTT 1740
 Db 1729 GAACACTATGTTGCTCAATGACCGCAACAAGCAATTTCTTCAACATGCAAGGCTT 1788
 Qy 1741 GATGTTGCAAGAACTTACCGGAGCTGTCAGTGTGTCGCGCAAGAACCAAGTGTGCT 1800
 Db 1789 GACCGTGTGAAGAACTTACCGGAGCTTGTGAGTGTGTAAGTAAAGTAAAGTAAAGTAAAG 1848
 Qy 1801 GAGTTGCTAAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1860
 Db 1849 GAGTTGCTAAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1908
 Qy 1861 GAGAAAGCTGTAATGAAGAAATGTTTGAAGCTGATGCAAGTAACTTGAACGCGCA 1920

Db 1909 GAGAAAGCCGAGATGAAGAAATGATACAGCTGATGAGACTCAAGATTGAACGCGCAA 1968
 Qy 1921 TTGAGATGATATCATCTCAATGAAGAAATCGAAATGTTGAATCTTACCGATACATT 1980
 Db 1969 TTGAGTGAATTTCTCTCAAGTAAACCTGTGTGAGAAATGAGAGCTGTACCGGTGATC 2028
 Qy 1981 TGCGACAGAAAGGTGCTTGTGTAACAGCTGCAATTTGATGAAGCTTTGATGACAGTT 2040
 Db 2029 AGTACACACAGGGGAGCTTTGTGACAGCTGCAAGTGTATGAGGATTTGGTTGACAGTG 2088
 Qy 2041 GTGAGCAATATGATCTGCGGTTTGCACAACTTGGCAACTGTAAACGTTGACAGCCGAG 2100
 Db 2089 GTTGAACCAATGACTTGTGGGTGTCACAACTTGGCCACATGCAATGTGTGCTGCTGAG 2148
 Qy 2101 ATATTTGATGAGGAATCTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATG 2160
 Db 2149 ATCATTTGATGAGGAATCTGTTTCAAGATGATGATGATGATGATGATGATGATGATGATG 2208
 Qy 2161 GACATCTGCTGCAATTTCTTGAAGAAAGTGAAGAAATCATCTCACTGGGATTAAGATC 2220
 Db 2209 GACCTCTTGTGATGTTCTTTGAGAACTGCAAGTTGAGGATCTCACTGGGACACATC 2268
 Qy 2221 TCCCAAGAGGCTTGAAGCAATGAGAGAGATGATCATGAAAGATTACTCGGAGAGA 2280
 Db 2269 TCTCAAGCTGTCTCCAACTGATGAAGAGATGACATGCGCAAAATTTACTCCCAAGAGG 2328
 Qy 2281 CTATTTGACCTGACAGAGTGTATGATTTCTGGAAGAGATGTTCCACTGAAACGCGCT 2340
 Db 2329 CTCTCAACCTGACAGAGTGTATGATTTCTGGAAGAGATGTTCTGAACTGAAAGCTCTC 2388
 Qy 2341 GAGAGTGTGTTACTTGAAGATGTTTATGCTTTAAGTACGTAAGCTGAGTGAATCA 2400
 Db 2389 GAGAGTGTGTTACTTGAAGATGTTTATGCTTTAAGTACGTAAGCTGAGTGAATCA 2448
 Qy 2401 GTTTCATTGCGAGAGAGATTAATTGAACCTG 2431
 Db 2449 GTGCCCTTCTGTTGATGAATGAAGATG 2479

RESULT 6
 MTR131943
 LOCUS 2728 bp mRNA linear PLN 09-APR-1999
 DEFINITION Medicago truncatula mRNA for sucrose synthase.
 ACCESSION AJ131943
 VERSION AJ131943.1 GI:4584689
 KEYWORDS sucrose synthase; sucS1 gene.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 REFERENCE 1
 Hohnjec,N., Becker,J.D., Puhler,A., Perlick,A.M. and Kuster,H.
 TITLE Genomic organization and expression properties of the MsSUC1 gene,
 which encodes a nodule-enhanced sucrose synthase in the model
 legume Medicago truncatula
 JOURNAL Mol. Gen. Genet. 261 (3), 514-522 (1999)
 MEDLINE 99254811
 PUBMED 10323232
 REFERENCE 2 (bases 1 to 2728)
 AUTHORS Hohnjec,N.
 TITLES Direct Submission
 JOURNAL Submitted (07-JAN-1999) Hohnjec,N., Lehrstuhl fuer Genetik,
 Universitaet Bielefeld, Postfach 100131, Bielefeld,
 Nordrhein-Westfalen, 33501, GERMANY
 FEATURES
 location/Qualifiers
 1..2728
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultiivar="Jemalong"
 /db_xref="taxon:3880"
 /tissue_type="root nodule"


```
gene
CDS
1. .2728
/gene="sucS1"
43. .2460
/gene="sucS1"
/EC_number="2.4.1.13"
/codon_start=1
/product="sucrose synthase"
/protein_id="CAB40794.1"
/db_xref="gi:4584690"
/db_xref="GOA:O9T0M6"
/db_xref="SPTREMBL:O9T0M6"
/translation="MATERLRTRVSLKRLDETLTANRETLTSLSPKAGKGLIHH
HOVIAEPEIPEDSRQKLTGAPGEVLRSTDEALVLEPPVALAPRPPGIREYRANV
NALVEMIQPAEFLKPKRELVDGSGANGFVLEDPPEPPYLPFPPTLANKSIGNGVPL
NRHLASLIFDKESLHPLLEFLRLHSYKTLMDRIQNDLSQHLRLAKEXLSTL
DPEYSEFEHFEQELGLERGMDTAERLVESIQLLDLDELAPOPCLETFLEFDRIPVY
FNVLISPHGFADDDVLYGPDGGOVVYILDQVRLSEMLSRIKKQGLDIPRILI
ITRLPRAVETTCQRLKRYGTGHECHILRRYPRPDYGIYARKWISREVPYLETYE
DVAHELAELQKFDLIVGNTSDQIVASLAKLGVYQCTTAHALERTKYPESDIYM
KKEEKKHFSQCFADLPANMHTDPIITSTOEIAGSDKVGQESHYATFLPELYEV
VHGIDVPDENIVSPGADQTIYPPYTESRRLTSFPYELIELYSVEVEEHIQVVK
DRNKPIETMARLDRVKNI TGLVEMYGNALRLVNLVVAAGRRKESDLEIAEM
KKWGLIETVYKNGOPRMISSQMRVNRGELYRVIDPKGAFVGPAYEARGLTVVA
MATGLPRTATNGPRAEIVHGKGFPHIDPHGRADLDIYEFKKYKNDPSHDKXS
OGGLQRIEKTMTTISORLLITLVGVGFWRHSNDRLESRRYLEMFTALKRKLAE
SVPLAVE"
```

ORIGIN

Query Match 59.4%; Score 1559; DB 8; Length 2728;
Best Local Similarity 78.0%; Pred. No. 0; Mismatches 531; Indels 3; Gaps 1;
Matches 1890; Conservative 0;

```
1 ATGGCTGAGCGTGTCTCATCTCGGCTCCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
|||||
43 ATGGCTACCGAAGCTTTGACTGTGTTTCATAGCCCTCAAGAGAGGCTTGATGAAACCTTA 102
|||||
61 CTGCTCAGAGAAAGAGATTTGGCCTTGTCTCTCAGAGATCGAGGCAAGAAAAGGA 120
|||||
103 ACTCTAATAGGAATGAAATTTGGCCCTTCTTCAAGGCTTGAAAGCAAGGAAAAGGA 162
|||||
121 ATTCTGCAACCACTCAAAATTATCTAGTTTGAAGTATCCCTGGAAGAAACAAGAA 180
|||||
163 ATTTTGCAACCACTCAAGTATGCTGAGTTTGAAGAAATTCCTGAAGATGTAAGACG 222
|||||
181 AAGCTGCTAATGCTGATTTTGTGAAGTATGAAAGCTAGTCAAGAGCATGTGTTG 240
|||||
223 AAGTTGACTGATGTCATTTGTAAGTTTGAAGTTCACACAGGAAGCAATGTTTTG 282
|||||
241 CTTCCATGAGTTCATCTGCTGTTGTCAGAGGCTGTGTTTGGAGTACATTAGAGTG 300
|||||
283 CCACCATGGGTTGCACTTGCTGTTGTCAGAGGCAAGGTATTTGGAGTATCTGAGAGTA 342
|||||
301 AATGTTACGCGCTTGTGTTGAGGAATCTCACTGTGTGAGTATCTCCAATTGAAGGA 360
|||||
343 AATGTCAGTCTGTGTGTGAAAATTGCAACCTGTGAGTTCTCAAAATTCAGAGGA 402
|||||
361 GAGCTTGTGATGAGAAAGTTCAAAATCTTGTGTTGAATGAAATTTGAGACCCCTTC 420
|||||
403 GAATTTGTGATGAGAGTCTAATGAGAACTTTGTGCTGAGTTGAGACTTTGAACATTT 462
|||||
421 AACTCATCATTTCCCGCCCACTCTTCAAAATTCATGTTAAATGTTGAGAGTTCTTA 480
|||||
463 ACTCATCTTTCCCTCGTCTTAATCTCAACAGTCAATTTGGAATGAGTGTGCAATTCCTT 522
|||||
481 AATGTACCTTTTGGGAAATGTTTCATGACAAAGAGACATGCAACCTTTGCTGAA 540
|||||
523 AATGCAACCTTTTGTCTTAACTCTTCATGACAAAGAGATTTACATTCACCTTTGAA 582
|||||
541 TTCTTCAGAGTTCATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAAC 600
|||||
583 TTTCTCAGACTTCACAGTCAAGAGGAAAGACATTTGATGTTGAATGACAGAAATTCAGAAC 642
|||||
601 TTGAATGCTTTCACATGTTTGTGAGAAAGACAGAGATCTTGTGTAACCTTACTCTCT 660
|||||
```

```
|||||
643 CCGATTTCTCTTCAACATGTTCTGAGGAAGCTGAAAGATATCTTAACACAAATGATCTT 702
|||||
661 GAGACACCATGTCGCCGAATTTGAAACCCGGTTCAGAGAAATCGGTTTGAAGAAGTTGG 720
|||||
703 GAAACACCATATCTCAGAAATTTGAAACAGAGTTCCAGAGATGGTTTGGAGAGAGTTGG 762
|||||
721 GGTGACACCGGAGAACCGGTGCTGAGATGATTCACATCTCTTTGATCTTCTTGAGGCA 780
|||||
763 GGAGACACCCGACAGCCGCTCTCGAGTCCATCCAACTTCTCTTGATCTTCTCGAGGCT 822
|||||
781 ACTGATCTCTGACACCTTTGAGAGTTCCTTGAGAAATCCCAATGAGTTCATAGTTGTG 840
|||||
823 CCCGACCTTTGACACCTTTGAGACTTCTCTTGATGAATCCCAATGAGTTCATAGTTGTG 882
|||||
841 ATTCTCACTCCCAACGAGATCTTCTGCTCAAGACAAATGTTTGGGGTATCCCGACACCGGT 900
|||||
883 ATCTTTCTCTCTCAATGTTACTTGTCTCAAGATGATGTCCTTGGAATACCTGATATCTGGA 942
|||||
901 GGGCAGGTTGTTTACATCTTTGATCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCGT 960
|||||
943 GGCCAGGTTGTTTACATCTTTGATCAAGTTCGTCCTTGGAAGAGCGAGATGCTCAGTCCG 1002
|||||
961 ATAAAGCAAGAGACTCAACATCAACCCCTCGAATCTCTCATTTACTAGACTTCTCTCT 1020
|||||
1003 ATTAAAGAAACAGGCTTGATATCATCTCTCGCATTTCTCATATCACCCGCTTCTCTCC 1062
|||||
1021 GATGCTGTGGAACCAACATGCGGTCAACGACTTGAGAAAGTATPCGGAACAGAGACTCG 1080
|||||
1063 GATGCACTCGGAACGACTTGTGCGCAACGACTTGAGAAAGTCTGAGAACTGAGACTTCG 1122
|||||
1081 GATATTTCTGAGTACCTTTCAGAAACAGAAAGGAAATGTTGCAAAATGATCTCAAGA 1140
|||||
1123 CACATCTTGAAGTCCCTTCAGAGATACGAAAGGAAATTTCCGCAATGGAATCTCAGCA 1182
|||||
1141 TTTGAAAAGTTCGCCATCTTGGAAAACCTACACAGAGATGTTGCTCATGAAATCTCC 1200
|||||
1183 TTTG--AAGTTCGGCCATATCTAGAAACCTTACACAGAGATGTTGCTCATGAGACTTGC 1239
|||||
1201 AAAGATTTGACGAGCAGCGCAGATCTGATCATCGAAACNACAGCGCAATATGCTC 1260
|||||
1240 AAAGATTTGACGAGCAACGATCTGATGTTGGAACATCAAGTATGAGAAACATTTGTT 1299
|||||
1261 GCTCTCTGCTCGACATAAATTAAGTTCACACAGTGCACCATGCGCATGCTTTGGAG 1320
|||||
1300 GCTCTTGTGTCACATAAATTAAGTTCACACAGTGCACCATGCGCATGCTTTGGAG 1359
|||||
1321 AAGACAAAATATCCAGATTCAGATATCTATTTGAGAAAGCTTGAAAGCAATACATTTTC 1380
|||||
1360 AAGACTAAGTACCCGCAATCCGATTTACTGGAAGAAATTCGAAAGAGATACACTTC 1419
|||||
1381 TCTTGCAATTTACAGCTGATCTTTTTCAGATGAAACATATCAGATTTTCAATCATGACAGT 1440
|||||
1420 TCTGCGCAATTTACCGTGAATCTTTTTCGCAATGAAACACACAGATTTTCAATCATGACAGT 1479
|||||
1441 ACTTTCAGAGAAATTTGACAGAGCAAGCACTGTTGTCATATCGAGAGCAACATGCT 1500
|||||
1480 ACTTTCAGAGAGATTTGTAAGAGCAAGCAAGGTTGACAGATATGAGAGTCAACATGCTC 1539
|||||
1501 TTGACTTCTCTGCTCTATCCGTTGTGTAACAGTATTCGATATGATGTTGATTTCCAAATTC 1560
|||||
1540 TTTACTCTTTCACAGACTTACCGGTGTCACAGATATGATATGATGTTGATTTCCAAATTC 1599
|||||
1561 AACATGTTTCCCTGAGTCTGATATGAGATATATCTTCCCTTACACCGAAGAGAAACGG 1620
|||||
1600 AACATGATATCTCAGAGAGTATCAGACATTTACTCTTTCACACCGAAGATGAGCCG 1659
|||||
1621 AGGTTGAAGATTTTCAATCTGAGATGAGAGACCTTCTTTCACACCAAGTTGAGATGAA 1680
|||||
1660 CGATTTGACATCTTCTTCACTGAAATGAGAGAGTCTTTCACAGTCAAGTTGAGAAATGAA 1719
|||||
1681 GAACATTAATGATGCTCATATGACCGGACAAAGCAATTTCTGTTCAACATGCAAGCTT 1740
|||||
```


Db 505 AACGTCACCTTTCTGCGCAACCTTCATGACAAAGAGTTGCATCCACTTTGGA 564
 Qy 541 TTCTCAGAGTCATTTGTCAAGGCGCAAGATGATGTTGAATGA CAGATTTCAAG 600
 Db 565 TTCTCAGAGTCATTTGTCAAGGCGCAAGATGATGTTGAATGA CAGATTTCAAG 624
 Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGATATCTTGGTACCTTACCTCT 660
 Db 625 CCTGATCTCTTCAACATGTTTGAAGAAAGAGATATCTTGGTACCTTACCTCT 684
 Qy 661 GAGACACCATGTGCGAAATTCGAAACCGGTTCCAGAAATCGGTTTGAAGAGGTTGG 720
 Db 685 GATACACCGTACTCGAAATTCGAAACAGGTTCCAGAGATGGTTTGAAGAGGTTGG 744
 Qy 721 GGTGACACCGGAGAACCGGCTCGAGATGATCCACTCTTTGGATCTTTGAGGCA 780
 Db 745 GGAACACCGGAGACCGGCTCGAGATGATCCACTCTTTGGATCTTTGAGGCT 804
 Qy 781 ACTGATCTTGCACCTTGAAGATTTCTTGGAGAAATCCCATGCTTCAATGTTGTC 840
 Db 805 CCTGATCTTGCACCTTGAAGATTTCTTGGAGAAATCCCATGCTTCAATGTTGTC 864
 Qy 841 ATTGCACTCCCAACGATTTGCTGCAAGACAAATGTTTGGGATTCGCAACCGGT 900
 Db 865 ATTCTTTCTCTCAGGTTACTTTGCTCAAGATGATGTTTGGATACCTGATACCGGT 924
 Qy 901 GGGCAGGTTGTTATCATCTTGAATCAAGTCCGAGCTTGAAGATGAGATCTCTCGT 960
 Db 925 GGTGAGGTTGTTATCATCTTGAATCAAGTCCGAGCTTGAAGATGAGATCTCTCGT 984
 Qy 961 ATAAAGCAACAGACTCAACATCAACCTCGAATCTCTATTAATTAATGATCTTCT 1020
 Db 985 ATTAAGCAACAGGCTGATATGATCTCTCGAATCTCTATTAATTAATGATCTTCT 1044
 Qy 1021 GATGCTGCGGAAACATGCGGTCACAGCTTGAAGAAATGATGCGAAACAGAGCACTG 1080
 Db 1045 GACGAGTCGGAACATGCTTGGCCACAGCTCGAAGGTTTATGAAACCGAGATTCG 1104
 Qy 1081 GATATTTCTTGAGTACCTTTCAGAACAGAAAGGAAATGTTGCAAAATGATCTCAGA 1140
 Db 1105 CACATTTCTTGAGTACCTTTCAGAGATCAGAAAGGAAATGTTGCAAAATGATCTC 1164
 Qy 1141 TTTGAAAAAGTCTGGCCATCTTGAAGAACTTACACAGAGATGTTGCTCATGAATCT 1200
 Db 1165 TTTG---AAGTCTGGCCATCTTGAAGAACTTACACAGAGATGTTGCTCATGA 1221
 Qy 1201 AAAGAGTTGACGCGACCGCATCTGATCATCGAAACNACAGGACCGCATATGCTC 1260
 Db 1222 AAAGAGTTGACGCGACCGCATCTGATCATCGAAACNACAGGATGAAACATGCT 1281
 Qy 1261 GCTCTCTTGTGCGACATTAATTAAGTGTCAACAGTGCACATCGCCCATGCTTGGAG 1320
 Db 1282 GCTCTCTTGTGCGACATTAATTAAGTGTCACTCAGTGTCTAATTTGCTCATGCA 1341
 Qy 1321 AAAGCAAAATATCCAGATTCAGATATCTATTTGAAGAGCTTGAAGACAAATACAT 1380
 Db 1342 AAAGCAAAATATCCAGATTCAGATATCTATTTGAAGAGCTTGAAGACAAATACAT 1401
 Qy 1381 TCTTGCAATTTACAGTATCTTTTGAAGAAACCATCAGATTTATCATATCAGCAGT 1440
 Db 1402 TCTTGCAATTTACAGTATCTTTTGAAGAAACCATCAGATTTATCATATCAGCAGT 1461
 Qy 1441 ACTTTCAGAAATATTCAGAGAACAGACATGTTGTCATTAAGAGGCACTGCT 1500
 Db 1462 ACTTTCAGAAATATTCAGAGAACAGACATGTTGTCATTAAGAGGCACTGCT 1521
 Qy 1501 TTTGCTCTTCTGCTCTTACCGTGTGTATCATGATGATGTTGATTCGCAATTC 1560
 Db 1522 TTTGCTCTTCTGCTCTTACCGTGTGTATCATGATGATGTTGATTCGCAATTC 1581
 Qy 1561 AAAGTGTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAAAGAGGCG 1620
 Db 1582 AAAGTGTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAAAGAGGCG 1641

Qy 1621 AGTTGAAGATTTCCATCTGAGATGAAGACCTTTTACACCAAGTTGAGATGCA 1680
 Db 1642 AGTTGAAGATTTCCATCTGAGATGAAGACCTTTTACACCAAGTTGAGATGCA 1701
 Qy 1681 GAACCTTATGTTGCTCAATGACCGGCAACGCAATTCGTTTCAATGCGGAGCTT 1740
 Db 1702 GAACCTTATGTTGCTCAATGACCGGCAACGCAATTCGTTTCAATGCGGAGCTT 1761
 Qy 1741 GATGCTCAAGAACTTAAACCGGACTCTGCGAGTGTGCGGCAAGAACCCAAAGTTG 1800
 Db 1762 GACGCTGTAAGAACTTAAACCGGACTCTGCGAGTGTGCGGCAAGAACCCAAAGTTG 1821
 Qy 1801 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATGAGCGGAAAGAAATTTAAAGTT 1860
 Db 1822 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATGAGCGGAAAGAAATTTAAAGTT 1881
 Qy 1861 GAGAGGCTGAATTAAGAAATGTTTGAAGTGTGAGCAAGTCACTTGAACGCGCAA 1920
 Db 1882 GAGAGGCTGAATTAAGAAATGTTTGAAGTGTGAGCAAGTCACTTGAACGCGCAA 1941
 Qy 1921 TTCAGATGATATCATCTCAATGAAACAGAAATCCGAATGTTGAACCTTACCGAT 1980
 Db 1942 TTCAGATGATATTCATCTCAATGAAACAGAAATCCGAATGTTGAACCTTACCGAT 2001
 Qy 1981 TCGCACGAAAGGCTCTTGTGACAGCTGCAATTTGATGAAGCTTTGATGACGTT 2040
 Db 2002 TGTGACCAAAAGAGGCTTGTGACAGCTGCTGCTGACAGGCTTTCGATCAACGTC 2061
 Qy 2041 GTGAGGCAATGACTGCGGTTTGGCAACCTTGGCAACCTGTAACGCTGACCGGAG 2100
 Db 2062 GTGAGGCAATGACTGCGGTTTGGCAACCTTGGCAACCTGTAACGCTGACCGGAG 2121
 Qy 2101 ATTAATGCTCAAGGAAATCTGTTTCAACATGATCTTCAACATGATGATCAAGCT 2160
 Db 2122 ATTAATGCTCAAGGAAATCTGTTTCAACATGATCTTCAACATGATGATCAAGCT 2181
 Qy 2161 GACATCTCTGCTGATTTCTTTGAAAGTGTGAAGAGATCAATCTCATGAGATGATC 2220
 Db 2182 GATCTCTGCTGATTTCTTTGAAAGTGTGAAGAGATCAATCTCATGAGATGATC 2241
 Qy 2221 TCCCAAGAGGCTTGAAGAAATGAGGAGAGATATCATGAGAAATTTACTCGGAGAA 2280
 Db 2242 TCCCAAGAGGCTTGAAGAAATGAGGAGAGATATCATGAGAAATTTACTCGGAGAA 2301
 Qy 2281 CTATGACCTCTGACAGAGTGTATGATTTCTGAGAGATGTTTCAACCTTGAACGCG 2340
 Db 2302 CTATGACCTCTGACAGAGTGTATGATTTCTGAGAGATGTTTCAACCTTGAACGCG 2361
 Qy 2341 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTAAGTGAAGCTGAAGCTGAATCA 2400
 Db 2362 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTAAGTGAAGCTGAAGCTGAATCA 2421
 Qy 2401 GTTCCATTTGACAGAGATG--AATGGAACCTGTTAAATAAATGAGGCGGTTTCTT 2458
 Db 2422 GTTCCATTTGACAGAGATG--AATGGAACCTGTTAAATAAATGAGGCGGTTTCTT 2481
 Qy 2459 GGAAGATTAATTTCTGTTTGAATTTCAATTTGAGAGAGCTCTTGTATTTCAATCT 2518
 Db 2482 ATCTGAGTGTGTTGTTTGAATTTTAAATTTGTAATTTGTAATTTGTAATTTG 2541
 Qy 2519 CTTTTCTTTTCTTTTTCGCGGCAATGTTTGAACATGAGGTTGTCGCGCTCAATT 2578
 Db 2542 TGTGATTAAGCTTTGATTAAGAAATGTCATGCTTTTCTTTTGAAGATTTGAA 2601
 Qy 2579 CCAATTAATATGATGATTTGTTTGAAGAAATTTGTTTGAAGAAATTTGTTTGA 2625
 Db 2602 TGTGATTAAGCTTTGATTTCTTTTGAAGAAATTTGTTTGAAGAAATTTGTTTGA 2648

RESULT 8
 AF049487
 LOCUS AF049487 2760 bp mRNA linear PLN 01-JUN-1998

DEFINITION Medicago sativa sucrose synthase mRNA, complete cds.
ACCESSION AF049487
VERSION AF049487.1 GI:3169543
KEYWORDS Medicago sativa
SOURCE Medicago sativa
ORGANISM Medicago sativa; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
REFERENCE 1 (bases 1 to 2760)
AUTHORS D. Lowell Robinson, Gian Trepp, Robert G. Gregerson, Scott N. Tuary, Robert Roeven, J. Stephen Gantt and Carroll P. Vance.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-1998) Agronomy and Plant Genetics, USDA-ARS-WMA and the University of Minnesota, 1991 Upper Buford Circle, 411 Borlaug Hall, St. Paul, MN 55108, USA
FEATURES Location/Qualifiers
source 1..2760
 /organism="Medicago sativa"
 /mol_type="mRNA"
 /cultivar="Saranac"
 /db_xref="taxon:3879"
 /tissue_type="nodule"
 62..2479
 /EC_number="2.4.1.13"
 /function="sucrose hydrolase"
 /codon_start=1
 /product="sucrose synthase"
 /protein_id="AACI7867.1"
 /db_xref="GI:3169544"
 /translation="MATERLRVSLKRLDELTANRNEILLSLRAKGGILLH
 HOVAEERIPESRQKLTGDAPEVARSQEAIVPWPVALVRPRGIEWIRUNV
 HALVENLQPAELFKELBELVDGSANGVLEDPPTASFPPTLANKSIGNVHL
 NRHLASLFDKESLHLELRLSHYKGLTLMNDRIQNDLSQHLRAEYLSLT
 DPEPYSEFHRFOEIGLERMGDTARVYLSIQLDLLEAPCTLESFLDRIPMV
 FNVLSPHGFADVDVIGYDPTGGOVYILDOVRALSEMLSIKOGSLIDRIIL
 ITRLPAVGTTCGRLEKVGTEHCHILRVPRDEGIVRKWISREWVPYETYE
 DVNHELKELSKRDLVGNVSDNIVASLLAHKLGCTOCTIAMALETKIKYPRDIYM
 KREKTHSCQPTADLPANMHTDPIITSTFOETAGSKDKVGQESTHATFLPELVN
 VHGIDVDPKENVISPGADQITYPYETSRRLTSFYBELLYSSVENEHLCVK
 DRNKPILFTMARLDVRKNITGLVEMYGKNAELRLVNLVVAGDRRESKDLIELAM
 KMYGLIETVLANGPRISSOMNRVANGELRYICDTKGAFOPAVYEAFLVVA
 MATGLPTEATLNGSPARIIVHKGSGPHIDPHGDRADLVPEPEKKAQPSHMDKS
 OGGRIEKKTYITISORLLITLGVGVFMKHSNLRSLRILEFYLKTRKLAB
 SVPLAVE"
ORIGIN
 Query Match 59.3%; Score 1556; DB 8; Length 2760;
 Best Local Similarity 77.9%; Pred. No. 0;
 Matches 1887; Conservative 0; Mismatches 531; Indels 3; Gaps 1;
 Oy 1 ATGGCTGAGCGTGTCTCACTCGCGTCCACAGTCTCCGTGAGCGTTTGATGAGACCTT 60
 Db 62 ATGGCTACAGAACGTTGACTCGTTCATAGTCTCAAGAGAGACTTGATGAACCTTA 121
 Oy 61 CTGGCTACAGGAACGAGATTTGGCGCTGCTCAAGATCGAGGGCAAGAAAAGGA 120
 Db 122 ACTGCTAATAGAGATGAATTTGGCCCTTCTTCAAGGCTTGAAGCAAAAGGAAAGGA 181
 Oy 121 ATTGTCGAACACATCAATATTCTAGAGTTTGAAGCTATCCCTGAAGAGAACAGAAAG 180
 Db 182 ATTTTGCAACCAATCAAGTATGCTGAGTTTGAAGAAATTCCTGAAGAGATGACAG 241
 Oy 181 AAGCTGCTAATGATGATTTTGAAGATGAAGGCTAGTCAAGAGCATGCTGTTG 240
 Db 242 AAGCTGATGATGATGATTTGAGATTGATGATTCACACAGAAAGCAATAGTTTG 301
 Oy 241 CTTCCAGAGGTTGACTGCTGCTGCTCCCAAGCGCTGCTGTTGGAGTACATTAAGCTG 300
 Db 302 CCACCAATGGGTTGACTGCTGCTGCTCCCAAGCGCTGATTTGGAGTATCTGAAGTG 361
 Oy 301 AATGTCAAGCGCTGTTGTTGAGAACTCACTGCTGAGTATCTCACTTCAAGAA 360

Db 362 AATGTGATGCTGCTGTTGTCGAAAAATTTGCAACCTGCTGAGTTTCAATTCAGAA 421
 Oy 361 GAGCTTGTGATGGAAGTTCAATGAAACTTGTGTTGGAATGGAATTTGAGCCCTTC 420
 Db 422 GAACCTTGTGATGGAAGTTGGAATGAAACTTGTGCTTGTGAGTGAACCAATTT 481
 Oy 421 AACTCATATTTCCCGCCCAACTCTTCAAAATCCATGGTATATGATGAGATTCCTA 480
 Db 482 ACTCATATTTCCCGCTGCTGCTCACTCAACAGTCAAGTGAATGAAATGGTGCATTCCT 541
 Oy 481 AATGTCACCTTTCCGCAAAATGTTTCCATGACAGAGAGATGACACCTTCTCTGAA 540
 Db 542 AATGCACTTTCTGCTGAATCTTTCATGACAGAGAGATTTACATCCACTTTTGA 601
 Oy 541 TTCTCAAGTCCATTTGTCACAGGCGCAAGAACATGATGTTGAATGACAGATTCAGAC 600
 Db 602 TTCTCAAGTCCATTCACAGGCGCAAGAACATGATGTTGAATGACAGAAATTCAAAC 661
 Oy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGATGATCTGTACCTTACCTCT 660
 Db 662 CCGATTTCTTCAACATGTTTGAAGAAAGCAGAGATGATCTTACACCAATGATCT 721
 Oy 661 GAGACACCATGTCGCAATTCGAACACCGGTTCCAGGAAATCGTTTGAAGAGTTGG 720
 Db 722 GAACACCATACTGAAATTTGAACACAGGTTCCAGAGATGTTGTTGAGAGAGCTGG 781
 Oy 721 GGTGACACCGAGAACGCTGCTGAGATGATCCACTCTTTTGAATCTTGTAGGCA 780
 Db 782 GAGACACCGCAGAGCGCTGCTGAGATGATCCACTCTTTGATCTTGTAGAGCT 841
 Oy 781 ACGATCTCTGACCCCTTGGAAGTTCTCTTGAGGAAATCCCAATGCTTCAATGTTGT 840
 Db 842 CCGACCTTCAACCTTGAAGTTCTTCTTGATGAAATCCCAATGCTTCAATGTTGT 901
 Oy 841 ATTCTCACTCCCAAGATCTTCTGCTCAAGACATGTTTGGGATATCCGACACCGGT 900
 Db 902 ATCTTCTCTCTATGATGTTTCTTGTGCTCAAGATGATGTTCTTGGATATCCGATGAT 961
 Oy 901 GGGCAGGTTGTTATCATCTTGAATCAAGTCCGACTTGGAAATGATGATGCTCTCGT 960
 Db 962 GGGCAGGTTGTTATCATCTTGAATCAAGTCCGACTTGGAAATGATGATGCTCTCGT 1021
 Oy 961 ATTAAGAAAGAGACCTCAATCAACCCCTGAAATCTCTATTAATTAAGATCTCTCT 1020
 Db 1022 ATTAAGAAAGAGACCTGATATCATCTCTGAAATCTCTATTAATTAAGATCTCTCT 1081
 Oy 1021 GATGCTGCGAAACAATGATGATCAACGATCGATCAAGATGATTAAGAAAGAGACCTCG 1080
 Db 1082 GATGCACTGCGAACTACTTGTGCGCAACGCTTGAAGAGATCTATGAACTGAGCATTCG 1141
 Oy 1081 GATATTTCTGAGTACCTTCAAGACAGAAAGGAAATTTGGAATGATGATCTCAAGA 1140
 Db 1142 CACATTTCTGAGTACCTTCAAGATGAGAAAGGAAATTTGGAATGATGATCTCAAGA 1201
 Oy 1141 TTGAAAAAGTCCGCAATCTTGAACCTTACACAGAGATGTTGCTCATGAAATCTCC 1200
 Db 1202 TTG--AACTGCGCATATCTAGAACTTACCTGAGATGTTGCTCATGAGCTTCC 1258
 Oy 1201 AAAGAGTTGACGCGACGCGAGATCTGATCATCGAAACACAGCGCAATATGCTC 1260
 Db 1259 AAAGAGTTGACGCGACGCGAGATCTGATCTGATGAGAACTACAGTGAATGAAACATGTT 1318
 Oy 1261 GCTCTCTGCTCGACATTAATTAAGTGTCAACAGTGCACATCGCCATGCTTTGGAG 1320
 Db 1319 GCTCTTGTGAGCACTAATTAATTAAGTGTCACTCAGGTGACATGCTGTGATGACCTCGAG 1378
 Oy 1321 AAGCAAAATATCAATCAATGATATGATATGAAAGAGCTTGAAGCAATCAATTC 1380
 Db 1379 AAGCAATGATCCGAGATGATATTTACTGGAAGAAATTCGAAGAGATGATCACTTC 1438
 Oy 1381 TCTTGCAATTTACAGTGAATCTTTTGAATGAAACCATGATTCATCATCACAGT 1440

D	b	259	AAACTAGCTGATGGGGACCTTTGGGGAAGTCTTGAGAGCCACTTCAGGAATCAATAGTTTGG	318
Q	y	241	CCTCCATGGGTTGCACTTGTCTGTTCGTCCAAAGCCCTGGTGTGTTGGAGTACATTAGATG	300
D	b	319	CCCCATGGGTGGCTCTTGCTGTTCGTCCAAAGGCCAAGGTGTGTGGGAATATATTAAAGTC	378
Q	y	301	AATGTTCAAGCCCTTGTGTGTGAGGAATCACTGTGTCTGATATCTCCACTTCAGGAA	360
D	b	379	AATGTCCATGACATGGTGTGTGGAAGAGCTTCAGATGTCTGATATTTTGCACTTGAAAGAA	438
Q	y	361	GAGCTGTGTGATGGAAGTTCAAATGGAAACTTTGTTTGGAAATTGAAATTTGAAGCCCTTC	420
D	b	439	GAACTGTGATGATGAAGTTGCAACGGTAACTTCGTTCTTGAGCTTGATTTTGAGCCCAATT	498
Q	y	421	AACTCATCATTTCCCCCGCCCAACTCTTTCAAATATCCATTGGTAAATGGTGGATTCCTTA	480
D	b	499	AATGCTTCCTTCTCTCGGCTCTACTCTCTCCAAAGTCCATTGGCCAAATGGGTTGAATTCCTC	558
Q	y	481	AATGTCACTCTTTGGGCAAAATGTGTCATGACAAAGAGACATGCACTCTTGTCTCGAA	540
D	b	559	AAACGTCACTCTCTCGCAAAAGCTTATTCATGATGTAAGAGATATGACGCCATTCTCGAAC	618
Q	y	541	TTTCTCAGAGTCCCAATGTCAAAAGGCAAGAAACATGATGTTGAATGACAAATTCAGAAC	600
D	b	619	TTTCTCAGAGTCCCTGCTACAAAGGAAAGACGATGATGTTGAATGATGAATTCAGAAC	678
Q	y	601	TTGAATGCTCTTCAACATGTTTGAAGAAACAGAGAGATCTTGGTACCCTACCTCTCT	660
D	b	679	CTGAATGCTCTTCCAGCATGTGCTGTGAGGAAGCAGAGATATCTAGCAACCTTGCACT	738
Q	y	661	GAGACACCATGTGCGCAATTCGAAACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG	720
D	b	739	GAAACACCCCTACTCCGAGTTTGCAAAACAAGTTCCAGAGAGATTGGCTTAGAGAGATGG	798
Q	y	721	GTGACACCCGAGAAACGGTGTCTGAGATGATCCACTCTTTTGGATCTTCTTGAAGCA	780
D	b	799	GGCAACACTGCGAGCGGTCTCGAAATGATCCAACTCTTTGGATCTCTCGAAGCC	858
Q	y	781	ACTGATCCCTTSCACCCCTTGAGAAGTCCCTTGGAGAAATCCCATGAGTTCATAGTTGTG	840
D	b	859	CTGATCTCTTGCACTTTCGAAAGTCTCTTGTGMAATCCCATGATGCTTTAATGTTGTCT	918
Q	y	841	ATTCTCACTCCCAACGGAATCTTCGCTCAAGACATGTTTGGGGTATCCCGAACCCGT	900
D	b	919	ATCTTTTCGCCCAAGGTTACTTTTGCAAAATATGCTTGGGTACCTCGATATCGGT	978
Q	y	901	GGGACGTTGTTTACATTTTGGATCAAGTCCGAGCTTTGGAAATGAGATGCTCTCCGT	960
D	b	979	GGCGAGGTGTTCATCTTGGACCAAGTGTGTCTTGGAGCATGAATCTTCAACGT	1038
Q	y	961	ATTAAGCAACAAGGACTCAATCATCACCCCTCGAATCCCATTTATCTAGACTCTTCCT	1020
D	b	1039	ATTAAGCAGCAAGAGTTTGAATATCTCTCGACTTCTCATTTACACCCGACTCTCTTCA	1098
Q	y	1021	GATCTGTGAGAACATGCGGTCAACGATTCAGAAAGTATACGGAACAGACACTCG	1080
D	b	1099	GATCAGTAGGAACAACCTGCACCAAGGTCTCGAGAAAGTATTTGGGACAGAGCACTCC	1158
Q	y	1081	GATATTTCTTGAATACCTTCAAGACAGAAAGGAAATTTGTCGAAATGATCTTCAAGA	1140
D	b	1159	CACATTTCTTCAGTTCCATTTCAAGAACAGAAAGGAATATGTCGAAAGTGTATCTCAGA	1218
Q	y	1141	TTTGAAGAAAGCTGGCCATACTTGGAAACCTACACAGAGATGTTGCTCATGAAATCTCC	1200
D	b	1219	TTTGT--AAGTCTGGCTTACTCTAGAGACTTACACCGAGATGTTTGCACAGAACTTACC	1275
Q	y	1201	AAAGAGTTGACGGCAACGACAGATCTGATCATTCGAAACNACAGCGACGCAATATGCTC	1260
D	b	1276	AAAGAGTTGACGGGAAACCGGATCTGATCATTTGGAACCTACAGAGACGGCAATATGTT	1335
Q	y	1261	GCTTCCTTGTCTCGACATAAATTAGGTGTCAACAGTGCACATGCCCCATGCTTTGGAG	1320

D	b	1336	GCCTCATTTGCTGGGACATATAAATCGGGAGTTACACAGTGTACTATCGCTCATGCTTGGAG	1395
O	y	1321	AAGCAAAATATTCAGATATTCAGATATCTATTTGGAAGAAGCTTGAAGACAATATCAATTC	1380
D	b	1396	AAAATCAATATCTGATTCAGATATTTATTTGGAAGAAGATTTGATGACAAGATATCACTTC	1455
O	y	1381	TCCTGGCCAAATTTACAGCTGATCTTTTGTGGCAATGAACCATACAGATTTTCAATCAATCACAGT	1440
D	b	1456	TCATCCCAATTCACAGCTGATCTTATTTGCGCATGAACCATACCGATTTTCATCATCAAGT	1515
O	y	1441	ACTTTCAGAGAAATTTGAGAGAGCAAGAGACACTGTGGTCAATACAGAGACCACTGCT	1500
D	b	1516	ACTTTCAGAGATTTGCTGGGACAAAGATACAGATTGACAAATATGAGAGTCAACAGCT	1575
O	y	1501	TTCACTCTTCCTGGTCTCTACCGTGTGTACATGATTCGATGTGTGATGCCAAATTC	1560
D	b	1576	TTCACTCTTCCTGGGCTCTACCGAGTGTGCCAGGAATGATGTGTGATGCCAAATTC	1635
O	y	1561	AACATTTGTTTCCCTGGTGTCTGATATGAGATATATCTTCCCTTACCCGAAAGAACCG	1620
D	b	1636	AACATTTGTTTCACTCGAGACAGATACGATACGATCTACCTCCGTACACAGAACCGAAG	1695
O	y	1621	AGGTTGAAGCATTTCCATCCCGAGATGGAAGACCTCTTTACACCAAAGTTGAGATGAA	1680
D	b	1696	AGGCTGACCTTCCTTCAACCTCGAATGAGAGCTTCTTATATGTAAGTTGAGATGAG	1755
O	y	1681	GAACACTTATGTGTGCTCAATGACCGGACAAAGCCAAATTCGTTCACATGCGCAAGCTT	1740
D	b	1756	GAACACTTGTGTGTATTTGAAGAAGTCGGTCCAGCGCATCATATTCACATATGCGCAATG	1815
O	y	1741	GATGTGTCAAGAACTTTAACCGGACTGTGAGTGTGCGGCAAGAACCCAAAGTTGCT	1800
D	b	1816	GATGTGTGAAGAACATATCACTGGACTGTGAGGTGATACGGCAAGAACAGCGCTTAGG	1875
O	y	1801	GAGTTGCTAACCTCGATTTGTATGTGTGATGATAGGGGAAAGAACTTAAAGATTTGAA	1860
D	b	1876	GAGTTGTGTAACCTAGTGTGTGTGCTGAGAACCGGAGAAAGGCTTAAAGACAATGAA	1935
O	y	1861	GAGAGGCTGAAATGAGAAATAATTTGAGCTGATCCACAAGTACAACTTGAACGGCCAA	1920
D	b	1936	GAGAGGCTGAGATGAGAAAGATGTATATCTTATTAAGAACCTTACAACTGGAATGGCCAA	1995
O	y	1921	TTCAATGATATCATCTCAATGACAGAAATCCGAAATGTTGAACTTTACCGATACAT	1980
D	b	1996	TTTATGATGATTTCTGCAAGATGAAACCGAGTAAAGAACGATGAAGTCTACCGATGAT	2055
O	y	1981	TGCGACACGAAAGGTGCTTTGTACACCTCGCATGTATATGAAGCTTTGGATATGACAGTT	2040
D	b	2056	GCAACACACGAAAGGAGCTTTTGTGCAACCTGCTGTGTACAGAGCTTTGCTTGGACAGTT	2115
O	y	2041	GTGAGAGCAATGATTTGCGGTTTGTGCAACATTCGCAACTGTAAACGCTGACACCGAG	2100
D	b	2116	GTGAGAGCCATGACCTGTGATTTCTTACCTTTTGCACATGCAATGAGAGTCCGTCTGAG	2175
O	y	2101	ATTATGTTCATGAGGAATCTGTGTTCAACATTGATCTTACCATGTGTATCAAGCTGT	2160
D	b	2176	ATTATCATTCACGGTAAATCTGGCTTCCATATCATATCATATACGATGATTCGCTGT	2235
O	y	2161	GACATATCTGCTGATTTCTTTGAAAATGTATGAAGAAATCCATCTTCACTGGGATTAATC	2220
D	b	2236	GAAATCTCTTGTGCACTTCTTCGAAAGAGCAAGGAAACCCATCTACATTTGGGATTAATC	2295
O	y	2221	TCCCAAGAGGCTTGAAGCAATGAGAGAGATACATGAAAGATTTACTCGAGAGAG	2280
D	b	2296	TCCCAAGAGGCTTCACGCTATCTACGAAAGATACATGCGAGATCTATTTGCGAAGG	2355
O	y	2281	CTATTTGACCTTGACAGAGTGTATGATTTTGGAAAGCATGTTCATCCTTGAACGCTG	2340
D	b	2356	CTGCTTAACTAATCATGAGTTTACGATTTACGATTTGGAAGCATGTATTCATCCTTGAACGCTG	2415
O	y	2341	GAGAGTGTGTATCTTTGAGATTTTATATGCTTTTAAATGACCGTAACTGCTGCTGATATCA	2400
D	b	2416	GAGAGCGCGTATATTTGAAATGTTTTACGCTTCAAGATACCGCAAGCTGCGCTGATTTCT	2475

FEATURES	Db	2401	2476
LOCUS	AF079851	2401	2476
DEFINITION	Plasm sativum nodule-enhanced sucrose synthase (nss) mRNA,	GTTCATTGCAGCAGAGTAATGGA	GTGCCACAGCCTGTGGATGATGAGCA
ACCESSION	AF079851	2401	2476
VERSION	AF079851	2401	2476
KEYWORDS	complete cds.	2401	2476
SOURCE	AF079851.1 GI:3377763	2401	2476
ORGANISM	Plum sativum (pea)	2401	2476
	Plum sativum	2401	2476
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Plum.	2401	2476
	1 (bases 1 to 2749)	2401	2476
	Fedorova,M., Tikhonovich,I.A. and Vance,C.P.	2401	2476
	Expression of C-assimilating enzymes in pea (Plum sativum L.) root nodules. In situ localization in effective nodules	2401	2476
	Plant Cell Environ. 22 (10), 1249-1262 (1999)	2401	2476
	2 (bases 1 to 2749)	2401	2476
	Fedorova,M., Tikhonovich,I.A. and Vance,C.P.	2401	2476
	Direct Submision	2401	2476
	Submitted (23-JUN-1998) Agronomy and Plant Genetics, University of Minnesota, 1991 Upper Buford Circle, St. Paul, MN 55108, USA	2401	2476
	Location/Qualifiers	2401	2476
	1..2749	2401	2476

ORIGIN	
Query Match	59.0%; Score 1548.6; DB 8; Length 2749;
Best Local Similarity	77.8%; Pred. No. 0; Mismatches 535; Indels 3; Gaps 1;
Matches 1882; Conservative	0;
QY	1 ATGGCTGAGCGTGTCTTCACTCGCGCTCCACAGCTCCCTGAGCGTTTGATGAGACCCCTT 60
DB	67 ATGGCTACTGATCGATTGACTCGTGTTCATATGCTCCGAGAGAGGCTTGATGAAACCTTG 126
QY	61 CTTCCTCACAGGAGACGATTTTGGCCTTGTCTTCAGATTCGAGGCGAAAGGAAAAAGGA 120

Db	127	ACTGCTAATGGAATGAAATTTAGCTCTTCTTCAAGAGTTGAAGCAAGGAAAGGA	186
Qy	121	ATTCTGCAACACCATCAAAATTTATTTAGAGTTGGAAGCTATCCCTGAAGAGACAGAAAG	180
Db	187	ATTTTGGCAACACATCAAGATGATTTGCTGATTTTGGAAATAATTCCTGAAGGAATGACAG	246
Qy	181	AAGCTGCTAATGAGTGCATTTTTTGAAGTAAATGGAAGCTATGTCAGAAAGGATGCTGTG	240
Db	247	AAGCTGACTAATGAGTGCATTTGGTGAAGTTCTAGATTCACACAGAAAGCTATAGTTTGG	306
Qy	241	CTTCATAGGGTTGCATCTGCTGTTCGTCAAGGCTGTGTGGAGTACATTAGAGTG	300
Db	307	CCACCAATGGGTGTGCTGTGCTTTCGTCCAAAGGCAGAGTGTCTGGAGTATCTGAGAGTG	366
Qy	301	AATGTTCAAGCCCTTGTGTGTTGAGAACTCACTGTCTGATATCTCCACTTCAAGAA	360
Db	367	AATGTGCATCTCTTGTGTGTTGAAATTTTCCAACTGTGATGTTCTTCAAAATTCAGAA	426
Qy	361	GAGCTGTATGGAAGATTCAAATGAAATCTTTGTTTGGAAATGATTTGATTTGAGCCCTTC	420
Db	427	GAACTGTATGGAAGATGCTAATGTAATGTAACTTTGTGCTTGAATTTGACTTTGAACATTT	486
Qy	421	AACTCATCATTTCCCCGCCCACTCTTCAAAAATCCATTTGTAATGTTGAGATTCCTA	480
Db	487	ACGCAATCTTTCCCTCGTCTTAATCTTCACAAATTCATTGGAATGTTGACTTTCTC	546
Qy	481	AATGTCACCTTTTGGCAAAATTTGTTTCATGACAGAGAGCATGCACCTTTGCTCGAA	540
Db	547	AAACGTACCTTTCTGCGCAAACTTTCATGACAGAGAGATTTGTCATCCACTTTTGAA	606
Qy	541	TTTCTCAAGTCCATTTGTCAAGGGGCAABAACATGATGTTGAATGACAAATTCABAAC	600
Db	607	TTTTCAGACTTCACAGCTACAAAGGGGAAACATTTGTTGAATGACAAATTCABAAC	666
Qy	601	TTGATGCTCTTCAACAATGTTTGGAGAAAGCAGAGAGATCTTGATACCTTACTCT	660
Db	667	CTGATCTCTTCAACAATGTTTGGAGAAAGCTGAAGATCTGAGCACAAGTTGCTCCG	726
Qy	661	GAGACACCATGTGCGCAATTCGAAACACCGGTTCCAGAAATCGGTTTGAAGAAGGTTGG	720
Db	727	GATACACCGTAACTCCGAATTCGAACACAGGTTCCAGAGATTTGGTTTGGAGAGGTTGG	786
Qy	721	GGTGAACCGGCAAAAGCGGTCTGAGATGATCAATCTCTTTGATATCTTGAAGGCA	780
Db	787	GGAGACCTGCAAGGCGGTCTGAGATTCATTCAGTTCTATATGATATCTTGAAGGCT	846
Qy	781	ACTGATCTCTGCAACCTTGAAGATTTCTTGGGAGAAATCCCAATGCTTCAATGTTGTG	840
Db	847	CTGATCTCTTGAACCTTGAAGATTTCTTGAAGAAATCCCAATGCTTCAATGTTGTG	906
Qy	841	ATTCTCACTCCCAAGGATATCTTGGCTCAAGACATGTTTGGGGATATCCCGACCGGT	900
Db	907	ATTCTTCTCTCATGAGTATCTTGGCTCAAGATGATGTTCTTGGGATATCCCGATACCGGT	966
Qy	901	GGGAGGTTTATTAACATTTGATGCAAGTCCGAGCTTTGGAGAAATGAGATGCTCCGT	960
Db	967	GGTCAAGGTTTATTAACATTTGATGCAAGTCCGAGCTTTGGAGAGTGAAGTCAATGCG	1026
Qy	961	ATTAAGCAACAAGACTCAACATCAACCCCTGCAATCTCTTATTAATTAAGACTTCTCT	1020
Db	1027	ATTAAGCAACAAGGCTTGAATTCGTTCTTCGCAATTCATTAATCACTCGTCTTCTCCA	1086
Qy	1021	GATGCTGTGCGAAACAATGCGGTCAAGCATTTGAGAAATATTCGGAACAAGACACTGC	1080
Db	1087	GACGCAAGTCCGAATCACTTGTGGCCAAAGCATCGAAGAAAGTCTATGGAACCGAGCATTGC	1146
Qy	1081	GATATCTTGAAGTACCTTCAAGACAGAAATGGGAATTTGTTGMAAATGATATCTCAAGA	1140
Db	1147	CACATCTTTCGAATTTCCCTTCAAGATTCAGAAAGGAATTTTTCGCAATGATATCTCCGT	1206
Qy	1141	TTTGAAGAAAGTCTGGGCATATCTTTGGAACCTTACACAGAGATGTTGCTCATGAAATCTCC	1200
Db	1207	TTTGC--AAATGCTGGCCATATCTAGAAACCTTACACGAGATGTTGCTCATGAGCTTGC	1263

QY 841 ATTCTCACTCCCAAGGATACCTTGCCTCAAGACAATGTTTGGGATATCCGACACCGGT 900
 DB 881 ATTCTTCTCTCTCAATGGTTACTTGTCTCAAGATGATGCTTGGGATACCTGATACGTGT 940
 QY 901 GGGGAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGATGAGATGCTCTCCGT 960
 DB 941 GGTGAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGATGAGATGCTCTCCGT 1000
 QY 961 ATAAAGCAAGAGACTCAACATCCCTCGAATCCCTCATTTACTAAGCTTCTCTCT 1020
 DB 1001 ATTAAAGCAAGAGCTTGGATGATGCTTCTGCAATCTCATTTACATCTGCTCTCTCT 1060
 QY 1021 GATCTGTGAGACAACATGCGGTCAACGACTTGAAGAAAGTATACGGAACAGACACTCG 1080
 DB 1061 GATCAGTGGGAACTACTGTGGCCAACGACTTGAAGAAAGTCTATGGAACCGACATGTGT 1120
 QY 1081 GATATTTCTGAGTACCTTCAAGAACAGAAAGGAAATTTGTCGAAAATGGATCTCAAGA 1140
 DB 1121 CACATTTCTGAGTTCCTTCAAGATCAGAAAGGAAATTTGTCGAAAGTGTATCTCAAGT 1180
 QY 1141 TTTGAAAAGCTGGCCATCTTGGAAACCTACACAGAGATGTGTGCTCATGAATCTCC 1200
 DB 1181 TTGG--AAGTCTGACATATCTAGAAACCTACACAGAGATGTGTGCTCATGAGCTTGC 1237
 QY 1201 AAAGAGTTGACGCGACGCGAGATCTGATCTCGAAAACNACAGCGACGCAATATGCTC 1260
 DB 1238 AAAGAGTTGACGCGAAAACGAGTCTGATGTGTGAAAACCTACAGTATGGAACATTTGT 1297
 QY 1261 GCTCTCTGCTGCGACATTAATAGGTGTCAACAGTGCACATGCGCATGCGCATGCTTTGAG 1320
 DB 1298 GCTCTTGTGTGACATTAATAGGTGTCAACAGTGCACATGCTCATGCTCATGCTTGAAG 1357
 QY 1321 AAGCAAAAATATCCAGATTCAGATATCTATTTGAAGAAAGCTTGAAGCAATATCATTTT 1380
 DB 1358 AAGACCAAGTATCTGAATGTGATTTTACGTGAAAATTTGAAGAAAGTATCATCTTC 1417
 QY 1381 TCTTGCCAAATTAAGCTGATCTTTTTCATGAACATATGATTTTATCATCATCAAGT 1440
 DB 1418 TCGTGCCAAATTAAGCTGATCTTTTTCGATGAACCAACGATTTTATCATCATCAAGT 1477
 QY 1441 ACTTTCAGAAATTCAGAGAACAGACACTGTTGTCAATACAGAGCCACATGCT 1500
 DB 1478 ACTTTCAGAGATTCGTGAAGACAGATATCTGTTGAACAGTATGAGAGTCAACATGCT 1537
 QY 1501 TTCACTCTTCTGCTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATGATG 1560
 DB 1538 TTCACTCTTCTGCTCTACCGTGTGTGTGACGATGATGATGATGATGATGATGATGATG 1597
 QY 1561 AACATTTGTTTCCCTGCTGATATGAGATATATCTTCCCTTACACCGAAGAGAGCGG 1620
 DB 1598 AACATTTGCTTCCAGAGAGCTGATGACATTTTACTTCCCTTACACCGAAGAGAGCGG 1657
 QY 1621 AGTTGAAGCAATTCATCTGAGATGAGACCTTCTTTACACCAAGTTGAGATGAA 1680
 DB 1658 AGTTGAAGCAATTCATCTGAGATGAGACCTTCTTTACAGACAGTGTGAAAATGAA 1717
 QY 1681 GAACACTATGTGTCTCAATGACCGCAACCAAGCAATTTGTTTCAAGATGCAAGGCTT 1740
 DB 1718 GAGCAGATATGTGTCTCAATGACCGCAACCAAGCAATTTGTTTCAAGATGCAAGGCTT 1777
 QY 1741 GATGCTGTCAGAACTTAACCGGACCTGTCAGATGTGTGCGGACAGAAACCAAGTTGCT 1800
 DB 1778 GATGCTGTCAGAACTTAACCGGACCTGTCAGATGTGTGCGGACAGAAACCAAGTTGCT 1837
 QY 1801 GAGTTGGCTAACCTGTAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1860
 DB 1838 GAGTTGGCTAACCTGTAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1897
 QY 1861 GAGAAAGCTGAATGAAGAAATGTTTGAAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1920
 DB 1898 GAGAAAGCTGAATGAAGAAATGTTTGAATGAGTGTGATGAGTGTGATGAGTGTGATGAG 1957

QY 1921 TTCAATGATATCATCTCAAAATGAACAGAAATCCGAAATGTTTAACTTTACCGATACAT 1980
 DB 1958 TTCAATGATATCTCAAAATGATGAAACCGGTGTGAAATGAGAGCTTTACCGGTATATC 2017
 QY 1981 TGGACACGAAAGAGTGTGTGTGACAGCCGTCATGATGATGATGATGATGATGATGATG 2040
 DB 2018 TGGACACGAAAGAGTGTGTGTGACAGCCGTCATGATGATGATGATGATGATGATGATG 2077
 QY 2041 GTGAGGCAATGACCTTGGGTTTGGCAACCTTGGCAACCTGTAACCGTGGACGACCGAG 2100
 DB 2078 GTTGAAGCCATGACCTGATGATTTACCAACATTTGCAACATGATGATGATGATGATG 2137
 QY 2101 ATTAATGTCATGAGAAATTTGTTTCAACATGATGATGATGATGATGATGATGATGATG 2160
 DB 2138 ATCAATGTCATGAGAAATTTGATTTCCATGATGATGATGATGATGATGATGATGATG 2197
 QY 2161 GACATCTGTGTGATTTCTTTGAAAAGTGTGAAAGATGATGATGATGATGATGATGATG 2220
 DB 2198 GATCTCTGTGTGATTTCTTGGAAAAGTTTAAAGCTGATGATGATGATGATGATGATG 2257
 QY 2221 TCCCAAGAGCTTTGAAACGAAATGAGAGAGATATATCATGAAAGATTTTACTGAGAGA 2280
 DB 2258 TCTCTGTGTGTGCTTCAAGCTATTTGAAGAGATGACATGGAATTTTACTCTCAGAG 2317
 QY 2281 CTATGACCTTGAACAGAGTGTATGATTTTGAAGAGATTTTCAACCTTGAACCGCT 2340
 DB 2318 CTCTTACACTCACTGTGTGTATGTTTCTGAAAACATGTTTCTTAACTTGCACCGCTC 2377
 QY 2341 GAGAGTGTGTGATACCTTGAAGTGTATGATTTTATGCTTATAGTGTGATGATGATGATG 2400
 DB 2378 GAGAGCGCGCTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 2437
 QY 2401 GTTCAATTTGCGAGAGATG 2420
 DB 2438 GTGCTCTACCTGTTGAGGA 2457

RESULT 13

BD236049

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PR

PC

CC

FH

FT

FT

FEATURES

source

1..2906

Location/Qualifiers

/organism="Eucalyptus grandis"

/mol_type="genomic DNA"

PAT 17-JUL-2003

linear

DNA

2906 bp

Materials and method for modification of plant cell wall

polysaccharides.

BD236049

BD236049.1

GI:33045819

JP 2002527056-A/71.

Eucalyptus grandis

Eucalyptus grandis

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; Myrtales; Myrtaceae; Eucalyptus.

1 (bases 1 to 2906)

Blokberg, L.N.

Materials and method for modification of plant cell wall

Patent: JP 2002527056-A 71 27-AUG-2002

GENESTS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE

FORESTS LTD

OS Eucalyptus grandis (flooded gum)

PN JP 2002527056-A/71

PD 27-AUG-2002

PR 08-OCT-1999 JP 2000575985

PC 13-OCT-1998 US 09/170862, 11-AUG-1999 US 60/148426 P1

CC LONARD NATHAN BLOKSBERG

PC C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N9/12, C12N15/00, C12N5/00

Materials and method for modification of plant cell wall

polysaccharides

FH Key

FT source

FT 1..2906

Location/Qualifiers

/organism="Eucalyptus grandis"

/mol_type="genomic DNA"

ORIGIN

/db_xref="taxon:71139"

Query Match 57.1%; Score 1499.4; DB 6; Length 2906;
Beech Local Similarity 76.5%; Fred. No. 0;
Matches 1852; Conservative 0; Mismatches 567; Indels 3; Gaps 1;

1 ATGGCTGACGCTGCTCACTCGAGTCCGATGAGCGTTGGATGAGACCTT 60
233 ATGGCTGATCCGATGCTGATGAGCAAGCAAGCTTCCGAGCGTTTGAAGAGCCCTC 292
61 CTGCTGACAGAAAGAGATTTGGCTCTGCAAGATGAGGCGAAAGAAAGAA 120
293 TCTGCTGACCGCAAGATGATGAGCTTCTTCAAGGTTGAAGCCAAAGGCAAGGC 352
121 ATTCGCAACACCATCAATTAATTTCTAGAGTTGAAGCTATCCCTGAAGAAACAAAG 180
353 ATCTTGACGCGCCACACAGATTTTGTCTGAGGCTATCTGAGAGAGAGAGACCA 412
181 AAGCTGCTAATGAGTATTTTGAAGATTTGAAGGCTAGTCAAGGAGCGATGCTGTG 240
413 AAGCTTCTGATGAGGCTTGTGTAAGTCTCAATATCACTCAAGAAAGCAATTTGTG 472
241 CCTCATGAGGTTGCACTTGTGTTGTCMAAGGCTGTGTTTGGAGTACATTAAGTGT 300
473 CCTCATGAGGTTGCTCTGTTGTTGTCMAAGGCGGCGGTGGAGACATCGGTGTG 532
301 AATGTTACGCGCTTGTGTTGAGAACTCACTGTTGCTAGATCTTCACTTCAAGAA 360
533 AAGCTGCAAGGCTGTTGTTGAGCAATTTGAGGTTGCTGAGATCTTCACTTCAAGAA 592
361 GAGCTGTTGAGTGAAGTTCAATGGAATTTGTTGGAATTTGGAATTTGAGCGCTT 420
593 GAGCTGCTGAGTGAAGCTTGAAGTGAATTTGCTTGAAGCTTGAAGCTTGAAGCTT 652
421 AACTCATATTTCCCGCCGCAACTTTCMAAATTCATTTGATTTGAGTGTGAGTTCTTA 480
653 AACTGCTTTTTCGCGCCGCACTTTTCAAGTTATTTGGAATTTGAGTGTGAGTTCTT 712
481 AATGTTACCTTTTGGGCAAAATTTGTTCAATGAAGAGAGATGCACTTGTGCTGAA 540
713 AATGCTCATCTCTCGCTGAGCTTTCATGACAAAGAAAGCTTGAAGCTTGTGTTAA 772
541 TTCTGCAAGTCAATTTGCAAGGCGCAAGAAAGATTTGAATTTGAATTTGAGAAC 600
773 TTCTGCAAGTCACTGCTACAGGGAAGAAAGATTTGATTTGCAAGTTCAGAAAT 832
601 TTGAATGCTTTTCAACATTTTGAAGAAAGCAAGAGATTTCTTGTATCCCTTACTT 660
833 GTGTTCTCTTCCAAACATGCTCTGAGAAAGCGAGAGATTTGACCTCGCTCAAAACC 892
661 GAGACACATGTCGCGAAATTTGAACAACGCTTCCAGAAATCGTTTGAAGAGTTGG 720
893 GAGACCGCTGATCTCCAGATTTGAGCAAAATTTCCAGAGATCGGCTGAGACGCGGCTG 952
721 GGTACACCGGAGAACGCTGCTGAGATTTGATTTCAACTCTTTTGGATCTTGAAGCA 780
953 GGTACACCGCTGAGCGCTGCTGAGATTTGATTTCAACTCTTTTGGATCTTGAAGCT 1012
781 ACTGATCTTCAACCTTGAAGATTTCTTGGAGAAATCCCAATGATTTCAATTTG 840
1013 CCGACCGCTGACCTCTGAGAAATTTGATTTGAGTTCCCAATGATTTCAATTTG 1072
841 ATTCCTACTCCCGAGATTTCTGCTCAAGAAATTTTGGGATTTCCGACACGCT 900
1073 ATCTATGCTCTCCCGAGATTTCTGCTCAAGAAATTTTGGGATTTCCGATTTCCGCT 1132
901 GGGAGGTTGTTTCACTTTGATCAAGTCCGAGTTTGAAGATTTGATTTGATTTGCT 960
1133 GGGAGGTTGTTTCACTTTGATCAAGTCCGAGTTTGAAGATTTGATTTGATTTGCT 1192
961 ATTAAGCAACAGATCTGATCACTCCCTGATCTCATTAATTTACTGACTTTCTTCT 1020

1193 ATTAAGCAACAGATCTGATTAATTTACTCTTCCGATTTCTCATTAATCACTCGCTTCTTCCA 1252
1021 GATGCTGCGGAGAAACAATGCGGTCAACGATTTGAGAAAGTATCGGAACAGAGACTCG 1080
1253 GACGCGGTGAGAACACTGTGCGCAGGCTTTGAGAAAGTTTGGACCGAGACTTC 1312
1081 GATATTTCTGAGTACCTTTCAGAACAGAAAGGAAATTTGTTGCAAAATGATCTCAAGA 1140
1313 CACTTTCTGCGCTTCCCTTCAAGAAATGAGAGAGTCTCCGCAATGATTTCCCGG 1372
1141 TTTGAAAAAGTCTGCGCATTTCTGAAAACCTACACAGAGATTTGCTCATGAATCTTC 1200
1373 TTCC--AGTGTGCGCTTATTTGAAAAGATACCTGAGATTTGCGCAGGAACTTGTCT 1429
1201 AAAGATTTGACCGGACCGCATCTGATCTGATCTGAAAACNACGCGCAATATGCTC 1260
1430 GGAAGTTGACGCGGCAAGCTGATCTGATCTGAAAACNACGCGCAATATGCTC 1489
1261 GCTCTCTTGTGCGACATTAATTTAGTGTCAACAGTGCACCATTCGCTTTTGGAG 1320
1490 GCTTCTTGTGAGCAATTAATTTAGTGTTCACAGTGTTCACATTTAGCCATGCTCTGAG 1549
1321 AAGCAAAATATTCAGATTCATATCTATTTGAAGAACTTTGAAGCAAAATACCATTTG 1380
1550 AAGACGAAGTACCGAGTCAACATATTACTGAAGAAATTTGAGAAATTAACCATTTG 1609
1381 TCTTGCCAAATTAACGCTATCTTTTGAATGAACCATTAACATTAATTAACCATTTG 1440
1610 TCTTGCCAAATTAACGCTATCTATCTATTCATGCAACCAACGCTTATTAATCAACGC 1669
1441 ACTTTCAGAAATTTGAGAAAGCAAGCACTGTTGCTCAATGAGAGCAACTGCT 1500
1670 ACTTTCAGAAATTTGAGAAAGCAAGCACTGTTGCTCAATGAGAGCAACTGCT 1729
1501 TTCACTCTTCTGCTCTACGCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1560
1730 TTCACTCTTCTGCTCTACGCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1789
1561 AACTTTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
1790 AACTTTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1849
1621 AGTTGAAAGCAATTTCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1680
1850 CGGTTGAATCTTCCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1909
1881 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
1910 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1969
1741 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
1970 GACGTTGCAAGAACTTGAACGCGCTTGTGAAGTATGAGCAAGAACTTCAAGATTTGAG 2029
1801 GAGTTGCTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
2030 GAATCTGCGCACTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2089
1861 GAGAGGCTGAATGAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1920
2090 GAGAGGCTGAATGAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 2149
1921 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
2150 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2209
1981 TGCAACAGAAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
2210 TGCAACAGAAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2269
2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2270 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2329

Qy	2101	ATTATTTGCCATGGGAAATCTGGTTTCAACATTGATCCTTACCAATGGTGAATCAAGCTGCT	2160
Db	2330	ATCATTTGGCATGGCAAAATCGGGCTACACATTGATCCTTACCAATGGTGAACAGCGCGCC	2389
Qy	2161	GACATACTCGTGGATTTCTTTGAAAAGGTATAGAAAAGATCCATCTCACTGGGATTAAGATC	2220
Db	2390	GAGCTTCTTTAGAACTTTCTTCAACAAGTGCAGAGATTGACAGTCCCACTGGGACGAGATC	2449
Qy	2221	TCCCAAGAGGCGTTGAAAACGAATAGAGAGAAATATACATGGAAATTTACTTCGAGAGA	2280
Db	2450	TCAAAGGGTGCATGCAAGACGAATTTGAAGAAAGTATACATGGAAATATATTCGAGAGG	2509
Qy	2281	CTATTGACCCCTGCACAGAGTGTATGGAATTTCTGGAAGATGTTTCCAACTTGAACCGCT	2340
Db	2510	CTGTGTAACCTGACCTGCGGTGTATGGCTTCTGGAAGACATGTGACTTAACTTGATCGGCGC	2569
Qy	2341	GAGAGTGGTGGTTACCTTGAAGATGTTTATGTCTCTTAAATACCGTAAGCTGGCTGAATCA	2400
Db	2570	GAGAGTCCGCGGTACTCTTGAAGATGTTCTATGTCCCTCAAGTATCGCCCATCTGGCACAGTCT	2629
Qy	2401	GTTCCATTGGCAGAGGAGTAAA	2422
Db	2630	GTTCTCCGGCTGTGAGGTAAA	2651

OY	61	CTTGTCTCAAGAAACGAGATTTTGGCCCTGTGCTCTCAAGATCTGAGGCCAAAGAGAAAGGA	120
Db	302	TCTGTCTACCGCAACGATATTTGGCCCTCTTTCAAGAGTTGAAGCCAAAGGCAAGGC	361
OY	121	ATTCTGCMAACCATCAATATTCTAGAGTTTGAAGCTATCCCTGAAGAGAAACAGAAAG	180
Db	362	ATCTTGAAGGCCCAACAGATTTTGGCTGAGTTTGAAGGCATCTGTGAAGAGACGAGAGA	421
OY	181	AACTTCGTAATGTCGATTTTTTGAAGTATTGAAGGCTAGTCAGAAAGCGATCGTGTG	240
Db	422	AAGCTTCGTGAATGGGGCCTTTGGAGAGTCTCTCAATCACTCAGAGAAAGCGATTTGTGCG	481
OY	241	CCTCCATGGGTGACCTGTGCTTCGTCCAAAGGCCCTGCTGTTTGGAGATCAATTAGAGTG	300
Db	482	CCTCAATGGGTTCCTCTTGTCTGTCTGTCCAAAGCCGGCCGCTGTGGAGACACATCCGTGTG	541
OY	301	AATGTCACGCCCTTGTGTGTGAAGAACTCACTGTGCTGAGTATCTCCACTTCAAGAA	360
Db	542	AACGTCCATGCGCTGTGTTCTTGACCAATATGGAAGTTGCTGAGTATCTCACTTCAAGAA	601
OY	361	GAGCTGTGTATGGAAGTTCAATATGAAGAACTTGTGTTTGAATTGGAATTTTGAAGCCCTTG	420
Db	602	GAGCTGTGTATGGAAGTTGAATGTAACCTTGTGCTTGAAGCTGACCTTGTAGGACATTC	661

RESULT	14
LOCUS	BD235993
DEFINITION	Materials and method for modification of plant cell wall polysaccharides.
ACCESSION	BD235993
VERSION	JP 2002527056-A/15.
KEYWORDS	Eucalyptus grandis
SOURCE	Eucalyptus grandis
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Myrtaceae; Eucalyptus.
REFERENCE	1 (bases 1 to 2913)
AUTHORS	Bloksberg L.N.
TITLE	Materials and method for modification of plant cell wall
JOURNAL	Patent: JP 2002527056-A 15-27-Aug-2002; GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE FORESTS LTD
COMMENT	OS Eucalyptus grandis (flooded gum) PN JP 2002527056-A/15 PD 27-Aug-2002 PF 08-Oct-1999 JP 2000575985 PR 13-Oct-1998 US 09/170862, 11-Aug-1999 US 60/148426 PT PC C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N9/12, C12N15/00, C12N5/00 CC Materials and method for modification of plant cell wall CC
FEATURES	FT key Location/Qualifiers FT source 1..2913 Location/Qualifiers 1..2913 /organism="Eucalyptus grandis" /mol_type="Genomic DNA" /db_xref="taxon:71139"
ORIGIN	
Query Match	57.1% Score 1497.8; DB 6; Length 2913;
Best Local Similarity	76.4%; Pred. No. 0;
Matches 1851, Conservative	0; Mismatches 568; Indels 3; Gaps 1,
OY	1 ATGCTGAGCGTGCCTCACTGCCGTCCACAGTCCTCCTGAGCGTTGGATGAGACCCTT 60
DB	242 ATGCTGATGCAGATTGACTGAAACCAACAGCCTTGCGAGCGCTTGGACGAGACCTC 301

Db	662	ACTGCTCTTTTCCGGCCCGAAGCTTTTCCAAAGTCATATGGCAATGGGCTGGAATTTCTC	721
OY	481	AATGTCACCTTTTCGGCAAAATTTGTTCCATGACAGAGAGCATGACCTTTGCTCGAA	540
Db	722	AATGGCATTCTCTCCGCTAAGCTCTTCCATGACAGAGAAAGCTTGACCTCTCTTGA	781
OY	541	TTCCCTCAGAGTCATTTGTGACAAAGGCGCAAACTGATGTTGAATGACAGAAATTCAGAAC	600
Db	782	TTCTTCCAAAGTCACCTGCTACAGGGGAGAAACTGATGTGAATGCCGAATCCAGAAAT	841
OY	601	TTGAATGCTCTTCAACATGTTTTGAGAAAGCAGAGAGATTCCTTGTAACCTTACTCTT	660
Db	842	GTGTTCTTCCCTCCAAACATGTCCTGAGAAAGCGAGAGATTCAGACCTCCCTCAAAACC	901
OY	661	GAGACACATGATGGCGAAATTCGACACCGGTTCCAGGAAATCCGGTTTGGAAAGAGTTGG	720
Db	902	GAGACCCCGTATCTCCAGTTTGACACAAAGTTTCCAGAGATCGGCTTCGACGGGGGTGG	961
OY	721	GGTGACACCGCAGAACCGGTGCTGAGATGATCCAACTCTTTTGGATCTTTGAGGCA	780
Db	962	GGTGACACCGGTGAGCGCGTCTCGAATGATTCAGCTCTGTGGATCTCTCTTGAAGCT	1021
OY	781	ACTGATCTTGACACCTTTGAGAAATTTCTTGGGAGAAATCCCAATGGTTCAAATGTTGTG	840
Db	1022	CCCAACCCGTGCACTTCGAGAAATTTTGGATATGGGTTCCCAATGGTTCCTCAACGTCTG	1081
OY	841	ATTCTCACTCCCAACGAGTACTTGTCTCAAGACATGTTTGGGGTATCCCGACACCGGT	900
Db	1082	ATCATGTCTCCCAACGAGTACTTGTGCTCAGAGACAGTCTTGATATCCGGAATACGGT	1141
OY	901	GGGCAAGTTGTTTACATCTTGGATCAAGTCCGAGCTTTGGAGATGAGATGCTCTCCGT	960
Db	1142	GGCCAGGTGTTTACATCTCGGATCAAGTTCGTGCCCTTAGAGAAAGAAATCTTCAACGC	1201
OY	961	ATTAAGACAAAGAGCTCAACATCAACCCCTCGAATCCATATTATTACAGCTCTTCTCT	1020
Db	1202	ATTAGACAAAGAGACTGGAATTTACTCTCTCGAATTTCTCATTTACATCCGGCTTCTTCCA	1261
OY	1021	GATCTGTCCGAAACAATCGCGTCAACGACTTGAAGAAATGATACGGAACAGAGCACTCG	1080
Db	1262	GACCCGGTTGGAAACAATGTGGCCACGCGCTTGAGAAAGTTTTTGGGACCGAATACCTCC	1321
OY	1081	GATATCTTCGAGATACCTTCAGAAACGAAAGGAAATTTGTGAAATGGAATCTCAGAA	1140
Db	1322	CACATTTCTTGGCTCTCCCTTCAGAAATGAGAGGGAATGTCGCCAAGATGATTTCCGGG	1381
OY	1141	TTTGAATAAGCTGCGCATACTTGGAAACCTTACACAGAGATGTTGCTCATGAATCTCC	1200

Db	1382	TTCC---	AGGTGTGCCTTATTTGGAAAGATACACTGAGATGTGCGAGGAACTTGCT	1438
Qy	1301	AAAGGTTGCA	CGGCACCGCACGCTGATCTGATCTATCGGAAACNAACAGCGACGACATATCGTC	1260
Db	1439	GGAGAGTTGCA	GGGGAGACGCTGATCTGATCTCGAAACCTACAGTATGGAACATTGTT	1498
Qy	1361	GCTTCCTTGCT	CGCACATAAATTAAGGTGTCAACAGTGCACATCGCCCATGCTTTGGAG	1320
Db	1499	GCTTCCTTGTTAG	CACTAAATTAAGGTGTTCACAGTGTATCAATAGGCCATGAGCCCTCGAG	1558
Qy	1321	AAAGCAAAAT	TATCCAGATTCCAGATATCTCATTTGGAAGAAAGCTTGAAGCAATACATTTCC	1380
Db	1559	AAGACGAAGTAC	CCAGAGTACAGACATTAATCTGAAAGAAATTTGAGAAATATACACTTC	1618
Qy	1381	TCTTGCCAA	TTTACAGTACTTTTTCGCAATGAAACCATATCAGATTTCATCATCACAGT	1440
Db	1619	TCTTGCCAGTT	CACTGCTGATCTCATATGCGCATGAACACACCGACTTATATACACAGC	1678
Qy	1441	ACTTTCAGAA	ATTCGAGAAAGACACATGTCATGTTGTCAATACGAGAGCCACACTGCT	1500
Db	1679	ACCTTCCAA	GAATGCTGAAAGCATACAGTGGGGCAGTATGAGAGTCAATGAAAC	1738
Qy	1501	TTCACTCTT	CCTGTGCTTACCGGTGTGTAATATGATATCGATGTGTGATTCCCAATTC	1560
Db	1739	TTCACTCTT	CTCGACCTCTACCGAGTGTTCACGGGATCGACGCTTCGACCCGAAAGTTC	1798
Qy	1561	AACATGTTT	CCCCGGTGTGATATGAGATATATCTCCCTTACACCGAAGAAAGCGG	1620
Db	1799	AACATTTGTT	CAACAGAGTGTCAATGACATCTACCTTGTTACACCGAACAGAGCGG	1858
Qy	1621	AGGTTGAAG	ATTTCCATCTTGAGATCGAAGACCTTCTTTCACACCAAGTTGAGATGAA	1680
Db	1859	CGGTTGAAT	CTCTTCCACCTGAGATCGAGAACTCTTCAAGGATGTTGAAGCAAG	1918
Qy	1681	GAACTTAT	TGTGTCTATGACCGCAACAGCCAAATTCGTGTTCAAATGCCAAGCTT	1740
Db	1919	GAACTTGT	GTGTGTTAAGATAGAAAGACCTATATTTTTCACCATGTGCAAGCTG	1978
Qy	1741	GATGTGT	CAAGAACTTAAACGGACCTGTGAGTGTGTCGGGCAAGAACCCAAAGTGTGCT	1800
Db	1979	GACGTGT	CAAGAACTTACACAGGCTGTGTAGTGTATGCGCAAGAACTCAAGTTGAGC	2038
Qy	1801	GAGTTGCT	AACTCTGTAGTGTATGAGTGTATAGCGCAAAAGGAACTTAAAGATTTGGAA	1860
Db	2039	GAACTCGG	CAACTGTGTGTGTGAGAGTACAGAGAAAGGATTTGGAAGCATTTGGAA	2098
Qy	1861	GAGAAAG	CTGAAATGAGAAATGTTTGAAGCTGATCGACAAATGACACTTGAACGCCAA	1920
Db	2099	GAGCAAGT	GTGAGATGAAGAAATATGACACCTCATCGAAAAGTACAGCTGAATGGCCAG	2158
Qy	1921	TTTCAGAT	GGATATCATCTCAATATGAACGAATTCGAAATGTGAACTTTACCGATACATT	1980
Db	2159	TTTCAGAT	GGATATTTCTCCCAATATGAACCGGGTAGAAATGAGAGCTCTACCGCTACATC	2218
Qy	1981	TGCGACA	AGAAAGGCTCTTGTACAGCGCTCATATTGTATGAAGCTTTGATGACAGTT	2040
Db	2219	TGTGACAG	AGAGAGCTTGCTTCAACCGGCTATCTATGAAGCTTTCCGGTTGACCGTG	2278
Qy	2041	GTGAGG	CAATGACTTGGGTTTTCGAACATTCGCAACCTGTAAACGTTGACACAGCCGAG	2100
Db	2279	GTTGAGG	CCATGACTGTGATGATTCGCAACTTTGCAATGTGAGGACACAGCTGAG	2338
Qy	2101	ATTATGT	CCAAATCTGTTTGAACATGATCTCTTACATGAGTCAAGCTGCT	2160
Db	2339	ATCATTTG	CAATGGCAATTCGGCTACCACTATGATCTTTCACATGAGTACAGCGCGCC	2398
Qy	2161	GACATAT	CTGTGCAATTTCTTTGAAAGATGTAAAGAAAGATCCATCTCACTGGATTAAGATC	2220
Db	2399	GAGCTCT	TGTATGACTTTTCAACAAGTGCAGAGATGTACAGTCCCATTTGGAGAGAGATC	2458
Qy	2221	TCCCAAG	AGGCTTGAACGAATAGAGGAAGTATATACATGAAGATTTACTCGAGAGA	2280

Dd	2459	TCAAGGGTGCATGACAGAAATTGAAGAAGATATCATGAAAAATATATCTGAGAGG	2518
Oy	2281	CTATTGACCCTGACAGAGATGTATGATTTCTGAGAGCATGTTCCAACTTGAAACGCCCT	2340
Dd	2519	CTGTTGAACCGTAACCTGCGGTATGAGCTTCCTGGAAGATGTACTAACCCTTGATCGGCGC	2578
Oy	2341	GAGAGTGTGCTTACCTTGAGATCCTTTTANGCTCTTAGTACCGTAAGCTGCGCTGAATCA	2400
Dd	2579	GAGAGTGGCGCGTACTTGAAATTTCTATAGCCCTCAAGATATGCGCCACATGCGCACAGTCT	2638
Oy	2401	GTTCCATTGGCAGAGAGATAA	2422
Dd	2639	GTTCTCCGCGCTGTGAGATAA	2660
 RESULT 15 BD262165			
LOCUS	BD262165	3103 bp	DNA Linear PAT 17-JUN-2003
DEFINITION	Composition and methods for the modification of gene expression.		
ACCESSION	BD262165		
VERSION	BD262165.1	GI:33071933	
KEYWORDS	JP 2002539834-A/57.		
SOURCE	Eucalyptus grandis		
ORGANISM	Eucalyptus grandis		
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Myrtaceae; Eucalyptus.		
AUTHORS	Perera,R., Rice,S.J. and Eagleton,C.K.		
TITLE	Composition and methods for the modification of gene expression		
JOURNAL	Patent: JP 2002539834-A 57 26-NOV-2002;		
COMMENT	GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE FORESTS INDUSTRIES LTD		
	OS Eucalyptus grandis (flooded gum)		
	PN JP 2002539834-A/57		
	PD 26-NOV-2002		
	PF 24-FEB-2000 JP 2000608755		
	PR 25-MAR-1999 US 09/276599,30-JUN-1999 US 60/146591 PT		
	RANJAN PERERA, STEPHEN J RICE, CLARE KATHERINE EAGLETON PC		
	CI2N15/09,A01H5/00,C07K14/415,CI2N5/10,CI2Q1/02,CI2Q1/68, PC		
	CI2N15/00,		
	PC CI2N5/00		
	CC Composition and methods for the modification of gene CC		
	expression		
	FH Key Location/Qualifiers		
	FT source 1..3103		
	FT Location/Qualifiers /organism='Eucalyptus grandis (flooded gum)'		
FEATURES			
source	1..3103		
	/organism="Eucalyptus grandis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:71139"		
ORIGIN			
	Query Match	57.1%; Score 1497.8;	DB 6; Length 3103;
	Best Local Similarity	76.4%; Pred. No. 0;	
	Matches 1851; Conservative	0; Mismatches 568; Indels	1; Gaps
Oy	1	ATGCGTAGCGGTCTCTCATCTCGCGCTCACAGTCTCCGTGAGCGTTGGATGAACCTT	60
Dd	447	ATGGCTATGTCATGTTGATCTCGAAGCCACAGCGCTTGGCGAGCGTTGGAGACAACCTTC	506
Oy	61	CTTGCTCACAGGAAGAGATTTTGAGCCCTGCTCTCAAGATCGAGGCGCAAGAAAAGGA	120
Dd	507	TCTGCTACCGCAACGATATTGTGTGCTCTCTTCAAGCGTTGAAGCCAAAGCGCAAAGCC	566
Oy	121	ATTCTGCAACCAATCAAAATTAATTCATGAGTTGAAGCTATCCTTGAAGACAAG	180
Dd	567	ATCTTGAGCGCCACACAGATTTTGTGATGTTGAGGCCATCTCTGAGAGACAGACA	626
Oy	181	AAGCTGCTAATGTGCAATTTTGAAGTATGAAGCTAGTCAAGAAAGCATGTGTTG	240
Dd	627	AAGCTTCTTAATGAGGCTTTGGTAAGTCTCTCAATCTCACTGAGAAAGCATGTGTG	686

```
QY 241 CTTCCATGGTTCGACTTGTGTCCTCAAGGCTGTGTTGGAGTACATTAGAGTG 300
Db CCTCCATGGTTCGACTTGTGTCCTCAAGGCTGTGTTGGAGTACATTAGAGTG 746
QY 301 AATGTTCAGGCTTTGTTGTGAGAGTCACTGTGTGATGATCTCCATTCAGAGAA 360
Db AATGTTCAGGCTTTGTTGTGAGAGTCACTGTGTGATGATCTCCATTCAGAGAA 806
QY 361 GAGCTGTGTGAGAGTTCAAATGAAACTTTGTTGGAATTGATTTGAGCCCTTC 420
Db GAGCTGTGTGAGAGTTCAAATGGAATTTGTTGCTTGAGCTTGACTTTGAGCCATTC 866
QY 421 AACTCATTTCCCGCCCAACTTTCAAAATTCATTGTTAATGTGTGAGTTCCTA 480
Db AACTCATTTCCCGCCCAACTTTCAAAATTCATTGGAATGTGTGAGTTCCTC 926
QY 481 AATGTCACTTTGGGCAAAATGTTCATGACAAAGAGACATGACCCCTTCTGAA 540
Db AATGTCACTTTGGGCAAAATGTTCATGACAAAGAGACATGACCCCTTCTGAA 986
QY 541 TTCTTCAGAGTTCATTGTCAAGGCAAGAAATGATGTTGAATGACAGAAATTCAGAAC 600
Db TTCTTCAGAGTTCATTGTCAAGGCAAGAAATGATGTTGAATGACAGAAATTCAGAAAT 1046
QY 601 TTGAATGCTTTCAACATGTTTGGAGAAAGCAAGAGATCTTGTAACCTTACTCT 660
Db TTGAATGCTTTCAACATGTTTGGAGAAAGCAAGAGATCTTGTAACCTTACTCT 1106
QY 661 GAGACACATGTGCGGAATTGGAAACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG 720
Db GAGACACATGTGCGGAATTGGAAACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG 1166
QY 721 GGTACACACCGGAGAACCGGCTCGAGATGATCCAACTCTTTTGSATCTTTGAGGCA 780
Db GGTACACACCGGAGAACCGGCTCGAGATGATCCAACTCTTTGAGTCTTTGAGGCT 1226
QY 781 ACTGATCTTGCACCTTTGAGAAAGTCTTGGAGAAATCCCATGTGTTCATGTTGTG 840
Db ACTGATCTTGCACCTTTGAGAAAGTCTTGGAGAAATCCCATGTGTTCATGTTGTG 1286
QY 841 ATTTCATCCCCCGAGATTTTGGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 900
Db ATTTCATCCCCCGAGATTTTGGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 1346
QY 901 GGGCAGGTTGTTTCACTTGTGATCAAGTCCGAGCTTTGAGAAATGAGATCTCTCGT 960
Db GGGCAGGTTGTTTCACTTGTGATCAAGTCTCTCGAGATTTCACTTCACTTCACTCCG 1406
QY 961 ATTAAGCAACAGGACTGACATCACTCCCTGGAATCTCATTTATTAAGTCTTCTCT 1020
Db ATTAAGCAACAGGACTGATATTAATCTCTCGATTTCACTTCACTTCACTTCTTCA 1466
QY 1021 GATGTGTGAGAAACAACTGTGGCTCAAGATTTGAGAAAGTATAGGAAACGACACTCG 1080
Db GATGTGTGAGAAACAACTGTGGCTCAAGATTTGAGAAAGTATAGGAAACGACACTCG 1526
QY 1081 GATAATTTCTGAGTACCTTTGAGAAACAGAAAGGAAATTTTGGAAATGAGATCTCAAGA 1140
Db GATAATTTCTGAGTACCTTTGAGAAATGAGAAAGGAAATTTTGGAAATGAGATTTCCCG 1586
QY 1141 TTTGAAAAAGTCTGGCCATTTGAGAAACCTACACAGAGATGTTGCTCATGAAATCTCC 1200
Db TTTG----AGTGTGGCCCTTATTTGAAAGATACACTGAGATGTGCGAGGAACTTGCT 1643
QY 1201 AAAAGATTGACGGGACCGGAGATGTATCATTCGAAACAAACGACGAGCAATATGCTC 1260
Db AAAAGATTGACGGGACCGGAGATGTATCATTCGAAACAAACGACGAGCAATATGCTT 1703
QY 1261 GCCTCTTGTCTGCAATTAATTAAGGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
Db GCCTCTTGTCTGCAATTAATTAAGGTGTCAACAGTGTCAATTAAGGTGTCAATTAAGGT 1763
QY 1321 AAGACAAATATTCAGATTTGAGATATCTATTTGGAAGAGCTTGAGACAAATACCATTTCC 1380
Db AAGACAAATATTCAGATTTGAGATATCTATTTGGAAGAGCTTGAGACAAATACCATTTCC 1764
QY 1381 TCTTGGCAATTTACAGGTGATCTTTTGTGCAATGAAACCATACAGATTTATCATCACAGT 1440
Db TCTTGGCAATTTACAGGTGATCTTTATCATATGACCAACCAAGCTTATATACCAAGC 1883
QY 1441 ACTTTCAGAAATTTGAGAGCAAGACACTGTTGGTCAATACGAGACCAACTGCT 1500
Db ACTTTCAGAAATTTGAGAGCAAGACACTGTTGGTCAATACGAGACCAACTGCT 1943
QY 1501 TTCACTTCTCTGCTCTACCTGTTGTACATGATATGATATGATGTTGATTCCTCAATTC 1560
Db TTCACTTCTCTGCTCTACCTGTTGTACATGATATGATATGATGTTGATTCCTCAATTC 2003
QY 1561 AACATGTTTCCCTGGTGTGATATGAGATATATCTTCCCTTACACGAGAGAGAGCGG 1620
Db AACATGTTTCCCTGGTGTGATATGAGATATATCTTCCCTTACACGAGAGAGAGCGG 2063
QY 1621 AGGTTGAGCAATTTCCATCTCGAGATGGAAGACCTTCTTTACACCAAGTTGAGATGAA 1680
Db AGGTTGAGCAATTTCCATCTCGAGATGGAAGACCTTCTTCAAGCAATTTGAGATGAA 2123
QY 1681 GAACACTTATGTGTCTCAATGACCGCAACAGCCAAATCTGTTCAATAGCCAGGCTT 1740
Db GAACACTTATGTGTGTGAAAGATTAAGAAAGCTTATTTTACATGAGGAGGCTG 2183
QY 1741 GATGTGTCAAGAACTTAACCGGACTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db GATGTGTCAAGAACTTAACCGGACTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAG 2243
QY 1801 GAGTTGCTTAACCTCGTATGTTGATGATGATAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Db GAGTTGCTTAACCTCGTATGTTGATGATGATAGGAGGAGGAGGAGGAGGAGGAGGAG 2303
QY 1861 GAGAGGCTGAATGAGAAATGTTTGAAGTGTGATGACAACTGATCACTTGAACGCGCA 1920
Db GAGAGGCTGAATGAGAAATGTTTGAAGTGTGATGACAACTGATCACTTGAACGCGCA 2363
QY 1921 TTTCAGATGATATCATCTCAATGAAAGAGATCGAAATGTTGAACCTTCCGATCATTT 1980
Db TTTCAGATGATATCTTCCAGATGAAACCGGATGAGAAATGAGAGCTTACCGCTTACATC 2423
QY 1981 TGGCACAAGAAAGGTCCTTTGTACAGCTGCAATGTATGAGAGCTTTGATGACAGTT 2040
Db TGGCACAAGAAAGGTCCTTTGTACAGCTGCAATGTATGAGAGCTTTGATGACAGTT 2483
QY 2041 GTGAGGCAATGACTTGGCTTGGCAACATTTGCAACTGTGAACGCTGACACGCGAG 2100
Db GTGAGGCAATGACTTGGCTTGGCAACATTTGCAACTGTGAACGCTGACACGCGAG 2543
QY 2101 ATTATGTGCATGGGAAATGTGTTTCAACATTTGATCTTACCATGATGATCAAGTGTCT 2160
Db ATTATGTGCATGGGAAATGTGTTTCAACATTTGATCTTACCATGATGATCAAGTGTCT 2603
QY 2161 GACATACCTGTCGATTTCTTTGAAAGTGTAAAGAGTCACTCACTGGGATTAAGATC 2220
Db GACATACCTGTCGATTTCTTTGAAAGTGTAAAGAGTCACTCACTGGGATTAAGATC 2663
QY 2221 TCCCAAGAGGCTTGAACGAATGAGAGAGATATCATGAAAGATTTACTGGAGAGA 2280
Db TCCCAAGAGGCTTGAACGAATGAGAGAGATATCATGAAAGATTTACTGGAGAGA 2723
QY 2281 CTATTTGACCTGACAGAGTATGATTTGGAAGGATTTTCCAACTTGAACGCGCT 2340
Db CTATTTGACCTGACAGAGTATGATTTGGAAGGATTTTCCAACTTGAACGCGCT 2783
QY 2341 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTAAGTACGTAAGCTGTGATATCA 2400
Db GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTAAGTACGTAAGCTGTGATATCA 2843
QY 2401 GTTCCATTGGCAGAGAGTAA 2422
```

Wed Jun 9 11:54:30 2004

us-10-003-405-1.rge

Page 25

Db 2844 GTCTCCGCGCTGTCGAGTAA 2865

Search completed: June 5, 2004, 05:44:35
Job time : 16241 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:49:54 ; Search time 22 Seconds
(without alignments)
1891.388 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247

Sequence: 1 MAERLTVHSLRERLDELTL.....EMFVLRKYLAEVPLAE 806

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3643.5	85.8	805	4	US-09-598-401C-77
2	2841	66.9	766	2	US-08-553-436A-8
3	1747	41.1	806	4	US-08-684-005-2
4	506	11.9	125	4	US-09-598-401C-76
5	494.5	11.6	720	4	US-09-394-272-14
6	437	10.3	1059	4	US-09-394-272-5
7	436	10.3	1068	2	US-08-429-054A-11
8	436	10.3	1068	2	US-08-718-777-7
9	436	10.3	1068	3	US-09-051-341-7
10	435	10.3	1068	4	US-09-394-272-8
11	435.5	10.3	1083	4	US-09-394-272-11
12	431	10.1	1056	4	US-09-394-272-1
13	421.5	9.9	1045	2	US-08-553-436A-6
14	421.5	9.9	1045	2	US-09-394-272-7
15	419	9.9	1054	1	US-08-356-354-4
16	419	9.9	1054	2	US-08-778-656-4
17	418.5	9.9	1081	4	US-09-394-272-6
18	413.5	9.7	1053	4	US-09-394-272-6
19	408	9.6	1059	4	US-09-697-367-16
20	407.5	9.6	963	4	US-09-394-272-12
21	407.5	9.6	963	4	US-09-394-272-13
22	407	9.6	1057	4	US-09-697-367-23
23	407	9.6	1057	4	US-09-394-272-2
24	406	9.6	1057	3	US-08-853-948B-2
25	400	9.4	1084	4	US-09-394-272-9
26	394	9.3	1054	4	US-09-394-272-3
27	387.5	9.1	1049	4	US-09-394-272-10

28	367.5	8.7	846	1	US-08-356-354-2	Sequence 2, Appl1
29	367.5	8.7	908	2	US-08-778-656-2	Sequence 2, Appl1
30	305	7.2	846	1	US-08-356-354-6	Sequence 6, Appl1
31	305	7.2	908	2	US-08-778-656-6	Sequence 6, Appl1
32	286.5	6.7	668	4	US-09-697-367-2	Sequence 2, Appl1
33	222.5	5.2	368	4	US-09-697-367-24	Sequence 24, Appl1
34	209.5	4.9	365	4	US-09-697-367-8	Sequence 8, Appl1
35	192.5	4.5	341	3	US-08-853-948B-4	Sequence 4, Appl1
36	189.5	4.5	343	3	US-08-853-948B-1	Sequence 3, Appl1
37	184.5	4.3	343	3	US-08-853-948B-10	Sequence 10, Appl1
38	182.5	4.3	358	4	US-09-697-367-20	Sequence 20, Appl1
39	182	4.3	348	3	US-08-853-948B-5	Sequence 5, Appl1
40	153.5	3.6	502	4	US-09-134-001C-4511	Sequence 4511, Ap
41	153	3.6	418	4	US-09-489-039A-8158	Sequence 8158, Ap
42	150.5	3.5	418	4	US-09-107-532A-6073	Sequence 6073, Ap
43	150	3.5	393	4	US-09-252-991A-19043	Sequence 19043, A
44	148.5	3.5	210	4	US-09-697-367-10	Sequence 10, Appl1
45	142.5	3.4	369	4	US-09-543-681A-7514	Sequence 7514, Ap

ALIGNMENTS

```
RESULT 1
US-09-598-401C-77
; Sequence 77, Application US/09598401C
; Patent No. 6596925
; GENERAL INFORMATION:
; APPLICANT: Perera, J. Ranjan
; APPLICANT: Bagleton, Clare
; APPLICANT: Rice, Stephen J.
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1036C2
; CURRENT FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: PCT/N200/00018
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 805
; ORGANISM: Eucalyptus grandis
US-09-598-401C-77
Query Match 85.8%; Score 3643.5; DB 4; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

QY 1 MAERLTVHSLRERLDELTLAHRREILALSRIRGKGILOHQOITLLEPAIPENRK 60
DB 1 MADRLTSHSLRERLDELTLAHRNDIVAFSLRVAKKGILQROQIAEPFALSEERRA 60
QY 61 KLANGAFPEVILKASOEAVLPPWVALAVRPPGWVEYIRVWVHALVVELTVAEYLAFKE 120
DB 61 KLDDGAFPEVILKSTOEALVSPWVALAVRPPGWVEHIRVWVHALVQLLEAELHKE 120
QY 121 EIVDSSNGNFVLEIDPEFPNSFPRLTSLKISIGNGVEFLNRHLSAKLPHDKESMPLLE 180
DB 121 EIVDSSNGNFVLEIDPEFPNSFPRLTSLKISIGNGVEFLNRHLSAKLPHDKESMPLLE 180
QY 121 FLRVCHGKMMNDRIQNALOHVLRKKEEYIGTLPPETPCAEFEHROEIGLEBGW 240
DB 121 FLRVCHGKMMNDRIQNALOHVLRKKEEYIGTLPPETPCAEFEHROEIGLEBGW 240
QY 181 FLVHCVGKMMNDRIQNALOHVLRKKEEYIGTLPPETPCAEFEHROEIGLEBGW 240
DB 181 FLVHCVGKMMNDRIQNALOHVLRKKEEYIGTLPPETPCAEFEHROEIGLEBGW 240
QY 241 GDTARVLEMIQLDLLEATDPCTLEKFLGRIPWVFNVLITPHGYFAQDNVIGYPTDG 300
DB 241 GDTARVLEMIQLDLLEATDPCTLEKFLGRIPWVFNVLITPHGYFAQDNVIGYPTDG 300
```

QY 301 GOVVYILDOVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCORLEKYGTEHS 360
 DB 301 GOVVYILDOVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCORLEKYGTEHS 360
 QY 361 DILKVPFTEKGIYKRWISREPKWPIYETTEDVAHISKEHCTPDLIIIGNXSDGNIV 420
 DB 361 DILKVPFTEKGIYKRWISREPKWPIYETTEDVAHISKEHCTPDLIIIGNXSDGNIV 420
 QY 421 ASLHAKKLVGTCTAHLEKTKYPSDIYWKLEDKYHSCQFTADLIAMNHTDFTTS 480
 DB 421 ASLHAKKLVGTCTAHLEKTKYPSDIYWKLEDKYHSCQFTADLIAMNHTDFTTS 480
 QY 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSGADMEIYFPYTEER 540
 DB 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSGADMEIYFPYTEER 540
 QY 541 RLKHPHEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
 DB 541 RLKHPHEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
 QY 601 ELANLVVVGGRKESKDLSEKAKMKFELIDKNLNGOFPMITSSOMNRIRNVELYRYI 660
 DB 601 ELANLVVVGGRKESKDLSEKAKMKFELIDKNLNGOFPMITSSOMNRIRNVELYRYI 660
 QY 661 CDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 720
 DB 661 CDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 720
 QY 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTGVGFMKHSNLER 780
 DB 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTGVGFMKHSNLER 780
 QY 781 ESRRLYEMFYALKYRKLAEVPLAE 806
 DB 781 ESRRLYEMFYALKYRKLAEVPLAE 806

RESULT 2

US-08-553-436A-8
 ; Sequence 8, Application US/08553436A
 ; Patent No. 5866790

; GENERAL INFORMATION:
 ; APPLICANT: HESSE, Holger
 ; APPLICANT: MULLER-ROBER, Bernd
 ; TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
 ; TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE
 ; TITLE OF INVENTION: CONCENTRATION
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ostrolenk, Faber, Garb & Soffen
 ; STREET: 1180 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY

; COUNTRY: US
 ; ZIP: 10036-8403
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/553,436A
 ; FILING DATE: 17-NOV-1995
 ; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/EP94/01671
 ; FILING DATE: 20-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DE P 4317596.1
 ; FILING DATE: 24-MAY-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meilman, Edward

; REGISTRATION NUMBER: 24,735
 ; REFERENCE/DOCKET NUMBER: P/951-117
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 382-0700
 ; TELEFAX: (212) 382-0888
 ; TELEX: 236925
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 766 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-553-436A-8

Query Match 66.9%; Score 2841; DB 2; Length 766;
 Best Local Similarity 72.4%; Pred. No. 6.5e-269;
 Matches 541; Conservative 90; Mismatches 112; Indels 4; Gaps 4;

QY 62 LANGAFPEVLKASGEAIVLPPWVALAVRPRGWEYIRVNVHALVBEITYAEYLHKEE 121
 DB 7 LSDGPFSEVLKSAQEAIVPPFVALAVRPRGWEYIRVNVSEINVEOLITVSEYLHKEE 66
 QY 122 LVDSGSGNFVLEIDPEPENSFPRLTSKISGVGEFLNRHLISAKLPHDKESHPLLEF 181
 DB 67 LVDSGADHVVLEIDPEPENSFPRLTSKISGVGEFLNRHLISAKLPHDKESHPLLEF 126
 QY 182 LRVHCHGKMMMLNDRIQNLALQHLVKAEEYLTLPETPCAEFEHRFOEIGLERGNG 241
 DB 127 LRVHCHGKMMMLNDRIQNLALQHLVKAEEYLTLPETPCAEFEHRFOEIGLERGNG 186
 QY 242 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVIIITPRGYRAQDNVLCPTDG 301
 DB 187 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVIIITPRGYRAQDNVLCPTDG 246
 QY 302 GOVVYILDOVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCORLEKYGTEHSD 361
 DB 247 QIVYILDOVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCORLEKYGTEHSD 306
 QY 362 ILRVPRTEKGIYKRWISREPKWPIYETTEDVAHISKEHCTPDLIIIGNXSDGNIV 421
 DB 307 ILRVPRTEKGIYKRWISREPKWPIYETTEDVAHISKEHCTPDLIIIGNXSDGNIV 365
 QY 422 SLHAKKLVGTCTAHLEKTKYPSDIYWKLEDKYHSCQFTADLIAMNHTDFTTS 481
 DB 366 SLHAKKLVGTCTAHLEKTKYPSDIYWKLEDKYHSCQFTADLIAMNHTDFTTS 425
 QY 482 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSGADMEIYFPYTEER 540
 DB 426 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSGADMEIYFPYTEER 485
 QY 541 RLKHPHEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
 DB 486 RLKHPHEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 545
 QY 601 ELANLVVVGGRKESKDLSEKAKMKFELIDKNLNGOFPMITSSOMNRIRNVELYRYI 660
 DB 546 ELANLVVVGGRKESKDLSEKAKMKFELIDKNLNGOFPMITSSOMNRIRNVELYRYI 605
 QY 660 ICDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 720
 DB 606 ICDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 664
 QY 720 ADILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTGVGFMKHSNLER 779
 DB 665 ADILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTGVGFMKHSNLER 724
 QY 780 ESRRLYEMFYALKYRKLAEVPLAE 806
 DB 725 ESRRLYEMFYALKYRKLAEVPLAE 751

RESULT 3

US-08-684-005-2


```
US-09-394-272-14
; Sequence 14, Application US/09394272
; Patent No. 6472588
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/09/394,272
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-09-394-272-14
```

```
Query Match      11.6%; Score 494.5; DB 4; Length 720;
Best Local Similarity 27.7%; Pred. No. 4,3e-39;
Matches 147; Conservative 108; Mismatches 185; Indels 91; Gaps 21;
```

```
QY 279 VVITPHGYAODNV-LGY-PDTGGQVYIILDQVRAL-ENEMLRKIQOGLNITRILIT 335
DB 8 ILLISVHGLIRGENMELGRDADTGGQTKYVELARALVKNPQARVD-----LL 56
QY 336 TRLLPDD-AVGTTCGGRLEKVGTEHSDILRVFRTREKGIVRKMIKRFKWPVLETYTED 394
DB 57 TRLIDPKVDADYAPRELII--GDRRQIVRLGCGEBEYIAK-----EMLDYLDNFA-D 107
QY 395 VAHEISKELHGTPTDILIGNKSDGNIVASLAKLGVCTTAAHL---EKTXYPDSDIYW 451
DB 108 HALDVLKQRPDLVDIHSYVADAGVGRILSHQDLGLVHTGSHSGRSKRTLLSGIKA 167
QY 452 KKLDEKTHFSCQFTADLFAMNHTDIIITSTFOELIGSKDTVGCQYSHAFTPLGLYRVH 511
DB 168 DEIBSRVMMARINIEBETLGSARAVITSTHOEIA---EQYAOY----- 208
QY 512 GIDVDP-KENIVSGADMEIYEP-----YTEEKRLKHf--HEIEIDLTYTKVNEBH 562
DB 209 -DYQPODMVLIPGTDLKFTYPPKGNEMETPIVOELORFRRP----- 251
QY 563 LCVLDRNKPILFTWRLDRVKNLTGLVWCGSKNPKLRELANLVVVGDRRKESKDLSE 621
DB 252 -----RKPIILASRPDRPKNIHKLIAAGSQSLQALQANLVVAQN-RDDIDDLQSG 303
QY 622 -KAENKMPFELIDKYNLNGQFRWISSQMRIRNV-ELYRYICDTYGAFVQPALYBAFGLT 679
DB 304 PREVLTDLLITIDRYDLGKVAV--PKOQAEADVVALFRLTALSGVFINPALTEPFGILT 361
QY 660 VVEANTCGLPATCNCGSPAEIIVHKGSGFNIDPHAGQADILVDFEKKCKDPSHMDK 739
DB 362 LLEBAACGVPVIAIEGGEVDIILKQCNQYLINPDEVDIADKLL---KVLNDQOQOQF 417
QY 740 ISOGGLKRIEKKYTWKIYSERLTLTGAVGFMKVSHNLERESRREYLEMFY 790
DB 418 LSESLGEGKRIYSPSHVESY--LEAINALQOQSVLKRSDLKRRRTLIY 466
```

RESULT 6
US-09-394-272-5

```
; Sequence 5, Application US/09394272
; Patent No. 6472588
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/09/394,272
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1059
; TYPE: PRT
; ORGANISM: Vicia faba
US-09-394-272-5
```

```
Query Match      10.3%; Score 437; DB 4; Length 1059;
Best Local Similarity 28.1%; Pred. No. 3,5e-33;
Matches 153; Conservative 99; Mismatches 206; Indels 86; Gaps 26;
```

```
QY 279 VVITPHGYAODNV-LGY-PDTGGQVYIILDQVRAL-ENEMLRKIQOGLNITRILIT 336
DB 169 IVLISVHGLIRGENMELGRDADTGGQTKYVELARALGSG-----MGVY---RVDLIT 218
QY 337 RLL--PDAVGTTCGGRLEKVGTEHSDIIVHKGSGFNIDPHAGQADILVDFEKKCKDPSHMDK 739
DB 219 KQVSSPD-VDMWSYCEPTEMLAPRNTDEFQDDMGSSGAYIIRLPF---GPRNKYIRK-E 272
QY 383 KWPVLETYTED-VAN--EISKEL-----HGT-PDLIGNKSDGNIVASLAKLGV 431
DB 273 ELWPIEIPVDGAMGHIIQMSKALGEQIGSHAVWPVALHGHVADAGDSALLSGALNVP 332
QY 432 QCTTAHLEKTK---YPSDIIYKKLEDKTHFSCQFTADLFAMNHTDIIITSTFOELIAG 487
DB 333 MIFGSHSGRDKLEBOLKQGRISTDEINSTYKIMRIEABELADGTEIVITSTROEIE- 391
QY 488 SKDTVOGYESHTAFTPLGLYRVHVG-----IDVDPKKNISPGADMEIYEPY---- 535
DB 392 -----EQRVLYNGPD-PYLEBKIRARIRNVSCGYRMFRRASVLPFGMEFHNIAPLDGDI 445
QY 536 -TEEKRLKHfPEIEIDLTYTKVNEBHLCVANDRNKPILFTWRLDRVKNLTGLVWCG 594
DB 446 ETEPEGILDHAPQDPRIW-----SEIMRFPSPRKPVLIALARPKKNIITLVVAFG 499
QY 595 KNPGLRELANLVVVGGRRK-ESKDLSEKAKMKMFELIDKYNLNGQFRWISSQMRIRN 653
DB 500 ECRPLRELANLTLMGNRDGIDEMSSSTSSVLSVLKILIDCYDYGVAV--PRGHHKQSD 557
QY 654 V-ELYRYICDTYGAFVQPALYBAFGLTVEAMTSGLPATCNCGSPAEIIVHKGSGFNID 712
DB 558 VPDILRAAKTKGVFINPAFLIEPFGILTLEBAAGVLPWATKGGPVDIHNVLNDGLIID 617
QY 713 PYHDAQADILVDFEKKCKDPSHMDKISOGGLKRIEKKYTW---KIYSERLTLTGAV 768
DB 618 PHDEKSIADALL---KLVSQKQMAKORQGLKNI-HLFEMPHCKTYLSKIATCKPRH 672
QY 769 GFWK 772
DB 673 PQWQ 676
```

RESULT 7
US-08-429-054A-11

```
; Sequence 11, Application US/08429054A
; Patent No. 5917126
; GENERAL INFORMATION:
; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,
; APPLICANT: JEAN, VOELKER, TONI; GERVAIS, MONICA
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE (SPS),
; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: BIERMAN AND MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,054A
FILING DATE: 26-APR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 842,337
FILING DATE: 20-March-1992
APPLICATION NUMBER: PCT/FR 91/00593
FILING DATE: 18-July-1991
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: French 90402094.9
FILING DATE: 20-July-1990
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Charles A. Muellerian
REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 146.1137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
US-08-429-054A-11

Query Match 10.3%; Score 436; DB 2; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;
Db 185 HCHKKMMMLNDRIQNLNA-----LQHVLRKAEVYL-----GTLPP-ETPCAE 226
Qy 97 HLAKKKQLELEGIQRISARRKEQVRRATEDLAEDLSGEKGDITIGELAPVETTKK 156
Db 227 FEHRPOEIGLERGMDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMFVNVVILTPHG 286
Qy 157 FQNRSDTLV---WSDDNK-----EKKL-----YIVLISVHG 185
Db 287 YFAQDNV-LGV-PDTGGQVVYILDQVRALENEMMLRIKQGLNITP---RLIITRL-- 339
Qy 186 LVRGEMELGRSDTGGQVKKVVELARA-----MSMMPGVYRVDLFTROYSS 232
Db 340 PD-----AVGTTCCGRLKGVYGTESHDILRVPRTEKGIYVKWISRFKWPY 387
Qy 233 PDVDMSYEPTEMLCAGSNDGEGESG---AYIVRIPCGRDKYLKK---BALMPY 284
Db 388 LETYED-VAH--EISKED---HGTPLD---IIGNKSDGNIVASLAHLKGLVQCTIA 436
Qy 285 LQEPVDAALAHILNMSKALGEOVNGRPVLPVIGHYADAGVALLSGALNVPMVYLG 344
Db 437 HALEKTK---YPSDIYWKKLEDKYHNSCOFTADLPANHTDPIITSTFOEIASGKUTV 492
Qy 345 HSLGNKKEQLLKQGRMSKEIDSTYKIMRRIEGELALDASELIVITSTROI---DEQW 401
Db 493 GQYESHTAFTLBGLYRVVHGIDV-----FDPKFNIVSGADMELYF 533
Qy 402 GLYD-----GFDVKLEKVLARARRGVSGHGRYMRMVVPIPGMDPSNVV 446
Db 534 PY-----TEEKRLKHFEPEIEDLLYTKVNEEHLCVLNDNRKPILEFTMPRLDRVX 584
Qy 447 VHEIDIDGQDVKDDIVGLEGASPKSMPIWAEV-----MRFLTNNHKMILALSPDPKK 501
Db 585 NLTGIVKCGKNPKKRELANLVVGGDRKESKDLLEKAE-----MKGMFELIDKYNLNG 639
Qy 502 NITTLVKAFGECRPLRELANLTLLINGNR---DDIDMSAGNASVLTTLVLKIDIKDYDYG 557
Db 640 QPRWISQGNRIKRVNELVRYICDTGAFVQPALYEAFGLTVEANTGCLPFPATONGPA 699

Db 558 SVAF-PKHNOADVPEIRLAKKKGVFINPALTVEPGLTILEAAAGLPIVATNGGPV 616
Qy 700 EIVHSGSGFNIDPYHGDQADILVDPEKKKPSHMDKISOGGLKRIEKTMYKIXSE 759
Db 617 DITNALNGLVDPHDQALADALL-----KLVAQKNLMQECRRNGLRNI-HLYSMEPCR 671
Qy 760 RLIT 763
Db 672 TYLT 675

RESULT 8

US-08-718-777-7
Sequence 7, Application US/08718777
Patent No. 5981852
GENERAL INFORMATION:
APPLICANT: Van Asche, C.
APPLICANT: Lando, D.
APPLICANT: Bruneau, J. M.
APPLICANT: Voelker, T.
APPLICANT: Gervais, M.
TITLE OF INVENTION: MODIFICATION OF SUCROSE
TITLE OF INVENTION: PHOSPHATE
TITLE OF INVENTION: SYNTHASE IN PLANTS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Barbara Rae-Venter
STREET: 260 Sheridan Avenue, Suite 440
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718, 777
FILING DATE: NOT YET ASSIGNED
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175, 471
FILING DATE: 27-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE.072.02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)328-4400
TELEFAX: (415)328-4477
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-777-7

Query Match 10.3%; Score 436; DB 2; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;
Db 185 HCHKKMMMLNDRIQNLNA-----LQHVLRKAEVYL-----GTLPP-ETPCAE 226
Qy 97 HLAKKKQLELEGIQRISARRKEQVRRATEDLAEDLSGEKGDITIGELAPVETTKK 156
Db 227 FEHRPOEIGLERGMDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMFVNVVILTPHG 286
Qy 157 FQNRSDTLV---WSDDNK-----EKKL-----YIVLISVHG 185
Db 287 YFAQDNV-LGV-PDTGGQVVYILDQVRALENEMMLRIKQGLNITP---RLIITRL-- 339

```

Db 186 LVGRNEMELGRDSTGGQVKKVVELARA-----MSMMPGVYRDLFTROYSS 232
Qy 340 PD-----AVGTCGOREKVVGTESDILRVPRTEKIVKRWISREKVPY 387
Db 233 PDVMSYGEPTMLCAGSNDGEGMSGG--AYIVRIPCGRDKYLK-----EALMPY 284
Qy 388 LETYED-VAH--EISKEI-----HGTPLD---ITGNXSDGNIVASLLAHKLGVTQCTIA 436
Db 285 LQEFVDGALAHILNMSKALGEQVNGRPVLPYVIGHYADAGDVAALLSGALNPMVLTG 344
Qy 437 HALEKTK-----YPSDIYWKLEDKYHSCFTADLPFAMNHTDFTITSTFOEIASKDTV 492
Db 345 HSLGNKKEQLLKQGRMSKEEIDSTYKIMRIEGBELADASELVITSTROEI---DEQW 401
Qy 493 GQYESHTAFTPLBGLRVVHGIDV-----FDPKNIVSPGADMEIYF 533
Db 402 GLYD-----GPDVLEKYLBARARGVSGHGRYPMRVVIPPGMDFSNV 446
Qy 534 PY-----TEERKRLKHPREIEDLLYTKVNEBHLCVLDRNKPILFTMRDLRVK 584
Db 447 VHEDIDGQDVKDIIVGLEGASPKSMPIWAEV-----MRFLTNPHKMILALSPPDK 501
Qy 585 NITGLVEMCGKNPRLRELANLVVGGDRKESKDLSEKAB-----MKMFELIDKYNLNG 639
Db 502 NITTLVKAFGECRPLRELANLTLMGNR-----DDIDMSAGNASVLTTLVKLIDKIDYLG 557
Qy 640 QFRWISSQMRIRNVELRYICDTKGAFVOPALYEAFLVVEAMTCGLPTFATCNGPA 699
Db 558 SVAF-PKHNOADVBEIRLAAMKGVFINPALVPEFGLTLEBAAGLPIVATKNGPV 616
Qy 700 EIIYHSGSGFNIDPYHGDQADILVDFEKCCKDPSHMDKISQGLKRIEKEYTWKIYSE 759
Db 617 DITNALNGLVDPHDQNAIDALL-----KLVAADKNLMQECRRNGLRNI-HLYSPBHC 671
Qy 760 RLIT 763
Db 672 TYLT 675

RESULT 9
US-09-051-341-7
; Sequence 7, Application US/09051341
; Patent No. 6124528
; GENERAL INFORMATION:
; APPLICANT: Shemmaker, C. K.
; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,341
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/17351
; FILING DATE: 25-OCT-1996
; APPLICATION NUMBER: US 08/549,016
; FILING DATE: 27-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,200
; FILING DATE: 12-JAN-1995

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter, Ph.D.,
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.110.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400
; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-051-341-7

Query Match 10.3%; Score 436; DB 3; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4,56-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;

Qy 185 HCHGKNMMLNDRIQNLNA-----LQHYLRKAEYL-----GTLPP-ETPCAE 226
Db 97 HIAKKQLELEGIQIRISARKKEQVREATEDLADLSGEKGDITGELAPVETTKK 156
Qy 227 FEHFQETIGERMGDTAEKVLQMLDLLEATDECTLEKPLGRIPMPVNVILPHG 286
Db 157 FQRFSDLTV-----WSDNKK-----EKKL-----YVLISVHG 185
Qy 287 YFAODNV-LGY-PPYGGQVYIILDOVRALENEMLRIKOGLNITP---RLITIRLL-- 339
Db 186 LVGRNEMELGRDSTGGQVKKVVELARA-----MSMMPGVYRDLFTROYSS 232
Qy 340 PD-----AVGTCGOREKVVGTESDILRVPRTEKIVKRWISREKVPY 387
Db 233 PDVMSYGEPTMLCAGSNDGEGMSGG--AYIVRIPCGRDKYLK-----EALMPY 284
Qy 388 LETYED-VAH--EISKEI-----HGTPLD---ITGNXSDGNIVASLLAHKLGVTQCTIA 436
Db 285 LQEFVDGALAHILNMSKALGEQVNGRPVLPYVIGHYADAGDVAALLSGALNPMVLTG 344
Qy 437 HALEKTK-----YPSDIYWKLEDKYHSCFTADLPFAMNHTDFTITSTFOEIASKDTV 492
Db 345 HSLGNKKEQLLKQGRMSKEEIDSTYKIMRIEGBELADASELVITSTROEI---DEQW 401
Qy 493 GQYESHTAFTPLBGLRVVHGIDV-----FDPKNIVSPGADMEIYF 533
Db 402 GLYD-----GPDVLEKYLBARARGVSGHGRYPMRVVIPPGMDFSNV 446
Qy 534 PY-----TEERKRLKHPREIEDLLYTKVNEBHLCVLDRNKPILFTMRDLRVK 584
Db 447 VHEDIDGQDVKDIIVGLEGASPKSMPIWAEV-----MRFLTNPHKMILALSPPDK 501
Qy 585 NITGLVEMCGKNPRLRELANLVVGGDRKESKDLSEKAB-----MKMFELIDKYNLNG 639
Db 502 NITTLVKAFGECRPLRELANLTLMGNR-----DDIDMSAGNASVLTTLVKLIDKIDYLG 557
Qy 640 QFRWISSQMRIRNVELRYICDTKGAFVOPALYEAFLVVEAMTCGLPTFATCNGPA 699
Db 558 SVAF-PKHNOADVBEIRLAAMKGVFINPALVPEFGLTLEBAAGLPIVATKNGPV 616
Qy 700 EIIYHSGSGFNIDPYHGDQADILVDFEKCCKDPSHMDKISQGLKRIEKEYTWKIYSE 759
Db 617 DITNALNGLVDPHDQNAIDALL-----KLVAADKNLMQECRRNGLRNI-HLYSPBHC 671
Qy 760 RLIT 763
Db 672 TYLT 675

RESULT 10
US-09-394-272-8
; Sequence 8, Application US/09394272
; Patent No. 6472588
; GENERAL INFORMATION:

```

APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 8
LENGTH: 1068
TYPE: PRT
ORGANISM: Zea mays
US-09-394-272-6

Query Match 10.3%; Score 436; DB 4; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;
185 HCHKGNMMLNDRIQNLNA-----LQHVLRKAEYL-----GTLPP-ETPCAE 226
197 HLRKKKOLEEGIORISARRKEQOVREATEDLAEDISEGKDDTIGELAPVETTKK 156
227 FEHFOEIGLERMGDTAERVLEMIQLDLLEATDPCTLEKFLGRIPWVFNVLITPHG 286
157 FGRNPSDLTV---WSDDNK-----EKKL-----YIVLISVHG 185
287 YFAQDNV-LGY-PTGGGOVYVILDOVRALENMLRIKQGLNITP---RIIITRL-- 339
186 LVRGEMELGRDSDTGGGVVVELARA-----MSMAGVVRVLDLFROYSS 232
340 PD-----AVGTTCGGRLEKVVTEHSDILRVPRTEKGIYKRWISREKWPY 387
233 PDVMSYGEPTMELCAGSNDGEGMGSGS---AYIVRIPCGRDYTKK-----BALWY 284
388 LETYED-VAH--EISKEP-----HGTPLD---IIGNSDGNIVASLAKHKLGVTOCTIA 436
285 LQEPVDGALAHILNLSKALGEOVNGRPVLPVIGHVADAGDVALLSGALNVMVLTG 344
437 HALEKTK---YPSDIYWKLEDKYHNSCOFTADLPAMNHTDPIITSTPOEIAKSOTV 492
345 HSLGRNKEQLKQGRMSKEEIDSTYKIMRRIEGEELADASELVITSTROEI---DEQM 401
493 GQYSEHTAFTLPGCLRVRVHGIDV-----FDDKENIVSGADMEIYF 533
402 GLYD-----GFDVKLEKVLRAPARRGVSGHGRYMPRMVITPGMDFSNV 446
534 PY-----TEKRLKHPREIEDLLTYTVENEHLCVANDRNKPLFTMPRLDRVK 584
447 VHEIDGSDGVKDDIVGLEGASPKSMPTMAEV---MRPLTNHKKMILALSRPDRK 501
585 NLTLGVEMCGKPKRLBELANLVVGGDRRKESKDLSEKAE---MKQFELIDRYNLNG 639
502 NITTLVKAFGECRPLRELANLTLTIMGNR---DDIDDSAGNASVLTIVLTKLIDKYDYG 557
640 QPRMSSQMRNRYNELRYICDTGCAFYQAPALYEAFTVYEAMTCGIPFATCNGGPA 699
588 SVAF-PKHNOADVETIRLAAMKGVFTNPALVPEPGLTLEAAHAGPIVATKNGGV 616
700 EIIIVGSGFNIDPYHQADILVDFEKKCKDDSHMDKISQGLKEIEEKYTKWISYE 759
617 DITNLNNGLVDPHDQNALADALL---KLVAADKNLMEQCRNRGLRNI-HLISMPERHC 671
QY 760 RLTT 763
DB 672 TYLT 675

RESULT 11
US-09-394-272-11
Sequence 11, Application US/09394272
Patent No. 6472588
GENERAL INFORMATION:

APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 11
LENGTH: 1083
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-394-272-11

Query Match 10.3%; Score 435.5; DB 4; Length 1083;
Best Local Similarity 24.2%; Pred. No. 5.1e-33;
Matches 188; Conservative 128; Mismatches 272; Indels 189; Gaps 36;
120 BELVDSSNGNPNVLELDEPFN-----SSPPTTYSK-----SIGNVGFPLNR--HL 164
44 QEKVFGDNAGDQHEKVSPIKYFVEEYVNSFDESCLKYTIKATNTRESRNLENI 103
165 SAKLFHDKESMHP-----LEPLRVCHKGNMMLNDRIQNLNALQHVLRKAE---E 213
104 CMRIWHLARKKKQIYWDGCVRLSKRIEREGRDAEDLSESEKDKNDSEKSE 163
214 YLGLT-PPETPCAEFEHFOEIGLERMG--DTERVLEMIQLDLLEATDPCTLEKFL 270
164 VTTLEPPRDMPIRSEMOI-----WSEDDKSRLNYILIR-----QVEIGF 207
271 GRIPVFNVVI-----LTP-----HGYFAQDNV-LGY-PTGGGOVYVILDOVRAL 314
208 SDFVFNMLVGLTWCLYLPCTFNCSYHGLVGRGEMELGSDSDTGGGVVVELARA 267
315 NEMLRIRKQGLNITPRLIITRLPD-AVGTTCGGRLEKVVTEHSD-----ILRVYF 367
268 N-----TEGVH---RVDLITRQISPEVDYSGEPVEMLSCEPSSGSIIRIPC 317
368 RTEKGIYKRWISREKWPYETYTE-----DVAEHISKEHLCG---TPDLIIGKXSD 416
318 ---GSRDKYIPK-ESLWPHIPEFVDGALNHIIVSIRASLGQVNGKPIWVPIHGVAD 372
417 GNIVASLALHKLGYTQCTIAHALEKTKY---PBSDIYWKLEDKYHNSCOFTADLPAMN 472
373 AGEVVAHLAALNPMVLTGSLGRNKEQLKQGRITREDIDRTYKIMRRIEAEQSLD 432
473 HTDFIITSTPOEIAKSOTVQYSEHTAFTLPGCLRVRVHGIDV----- 515
433 AAEMVVTSTROEID-----AQW-----GLY---DGRDILKELRLVRRRGVSC 474
516 --FDDKENIVSGADMEIYFPT-----BEKRLKHPREIEDLLTYTK 556
475 GRYPVRMVVITPPGMD---FSYVLTQDSQEPDGLKSLIGPDRNDIKKPVPIW----- 524
557 VENENHLCVANDRNKPLFTMPRLDRVKNLGLVEMCGKPKRLBELANLVVGGDRRKES 616
525 ---SEIMKPFSPNHPPTILASRPDHKKNVTLTVAFECOPPLRELANLVILGR--- 577
617 KDLER-----FAEMKMPFELIDRYNLNGQFWMISSQMRNRYNELRYICDTGCAFYQ 671
578 DDIEMRSSSVLMNLVAKLIDQYDLVYGOVAV-PKHNOADVETIRLAAMKGVFTNPAL 636
672 LYFAFGTLVPEAMTCGLFTPATCNGGPAEIIIVHSGSGFNIDPYHQADILVDFEKKCK 731
637 LVPEPGLTLEAALVYGLPIVATRNGSPVDIYKALNGLLVDPHDQNALADALL---KL 692
QY 732 KDSHMDKISQGLKEIEEKYTKWISYERLTLTGCVYFMVGVNLEBERSRVLEM 788
DB 693 ANKHLMAECRNGKLNIT-HRFSWBPBCHRNLYS-----HVECHRNRPHTSSLDI 739

RESULT 12

Db 398 D-PVLERKLRAMKKGVSVCYGRFMRMVYIPRGME-----FNHIVPHEGMD 443
QY 554 YTKVNEEH-----LCVLNDRNKPILFTMPRLDRVKNLTGLVWCGKPKLRE 601
Db 444 GETETEETSPDPPIWAEIRFSPKPKPMILALARPDPCKNITTLVKAEGECRPIRE 503
QY 602 LANLVVGGDRRK-BSKOLEEKAEMKMPELDKNINGQPMISSOMNIRNV-ELRY 659
Db 504 LANLVINGNRGIDEMSTSSVLSVTKLIDYDLYGQVAY--PKHKQADVPEIRL 561
QY 660 ICDTGAFAVQALYEAFLTVVEAMTCGLPTFATCGNGPAEIIYHKGSGFNIDPFGDOA 719
Db 562 AKKTGVFINPAFIEPFGTLIEAAHGLPMVATNGGVPDIQVLDNGLVDPHEQOSI 621
QY 720 ADILVDFEKKCKDPSHMDKISQGLKRIEKKYTMKISERLLT 763
Db 622 ATALL-----KLAVADKQLMTKCOQNGLNKI-HLYSPHESKTYLS 660

RESULT 14

US-09-394-272-7
Sequence 7, Application US/09394272
Patent No. 6472588
GENERAL INFORMATION:
APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1045
TYPE: PRT
ORGANISM: Beta vulgaris
US-09-394-272-7

Query Match 9.9%; Score 421.5; DB 4; Length 1045;
Best Local Similarity 28.1%; Pred. No. 1.1e-31;
Matches 164; Conservative 93; Mismatches 210; Indels 117; Gaps 28;
QY 241 GPTARVLEMLQILDLEATDPCTLEKLGRIPMVFNVIITPHGYFAQDNV-LGY-PD 298
Db 133 GDSRPRPRIN-SIDAMETWISOKEKCL-----YVLISLHGLIRGENMEIGRSD 184
QY 239 TGGGVVYILDQVRALENEMLLRIKQGLNITPRILITRLT--PDVGTGCGRLKXV-- 354
Db 185 TGGGVKVVVELARALGS-----MPGVY--RVDLTRQVSSPD-VDMVSGEPTKMLNP 233
QY 355 ---YGTESD-----ILRVPRTEKGIIVRKWISFEKVMPLYLETYEDVAHEI-- 399
Db 234 RDSNGFDDDDDEMGESSGAYIVRIPE---GPRDKYIAK-ELMWIYIPEFVUGALNHIVQ 288
QY 400 -SK---ELHGT-----PDIIIGNKSDGNIVASLALHKLGVTOCTTAHALEKTKYPD--- 446
Db 289 MSKVLGEQIGSGGETWPAVHGHVADAGDSALLSGGLNVPMLLTGHSIGRDKLEQLKQ 348
QY 447 -----SDIYWKLEKYNHSCQFTADLFAMNHTDPIITSTPOEINGKDTVGQYESHAF 501
Db 349 GRMSKDDI-----NNTYKIMRRIEAEELSDASEIVITSTROEIE-----EQMHLVDF 397
QY 502 TLPGH-----YRVHAGIDV---FDPKFNVISPGADMEIYFYTEEKRRLKHFPPIBDL 553
Db 398 D-PVLERKLRAMKKGVSVCYGRFMRMVYIPRGME-----FNHIVPHEGMD 443
QY 554 YTKVNEEH-----LCVLNDRNKPILFTMPRLDRVKNLTGLVWCGKPKLRE 601
Db 444 GETETEETSPDPPIWAEIRFSPKPKPMILALARPDPCKNITTLVKAEGECRPIRE 503
QY 602 LANLVVGGDRRK-BSKOLEEKAEMKMPELDKNINGQPMISSOMNIRNV-ELRY 659

Db 504 LANLVINGNRGIDEMSTSSVLSVTKLIDYDLYGQVAY--PKHKQADVPEIRL 561
QY 660 ICDTGAFAVQALYEAFLTVVEAMTCGLPTFATCGNGPAEIIYHKGSGFNIDPFGDOA 719
Db 562 AKKTGVFINPAFIEPFGTLIEAAHGLPMVATNGGVPDIQVLDNGLVDPHEQOSI 621
QY 720 ADILVDFEKKCKDPSHMDKISQGLKRIEKKYTMKISERLLT 763
Db 622 ATALL-----KLAVADKQLMTKCOQNGLNKI-HLYSPHESKTYLS 660

RESULT 15

US-08-356-354-4
Sequence 4, Application US/08356354
Patent No. 5767365
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,354
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weillman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1054 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-354-4

Query Match 9.9%; Score 419; DB 1; Length 1054;
Best Local Similarity 27.4%; Pred. No. 2e-31;
Matches 155; Conservative 93; Mismatches 207; Indels 110; Gaps 24;
QY 267 EKFLGRIMVFNV-----VILTPHGYFAQDNV-LGY-PDTGGQVYVIL 307
Db 139 ESTGRRLRISVSVEWMAWVGQGRKLYIVLISHLIRGENNELGRDSDTGQGVKVV 198
QY 308 DQVRALEN-----EMLRKQGLNITPRI-----LIIRLBDPAVGTTCGQPLE 352
Db 199 ELARLGSMPGVRYDLTRQVS-----SPVDWSYGEPTKMLTPISTDGAMTENGE--- 250
QY 353 KYVGTESDILRVPRTEKGIIVRKWISFEKVMPLYLETYEDVAHEI---SK---ELHGT 406

```
Db 251 ----SSGAYIIRIPF---GPREKTIPIK-EQLMPYIPFVGDALNHIIQMSKVLGEQIGS 301
Oy 407 -----PDLIIGNXSDGNIVASLAKLGVTOCTTAHALEKTYPPSDIYWKLEDK---- 457
Db 302 GYPWPVPAIHGHYADAGSALLSGALNVPMIFGHSIGROKLEQLLAQGRKSKDEINST 361
Oy 458 YHFSQFTADLFAMNHTDFIITSTFOEIAGSKDVGQYESHAFLLPGLYRVVHG---- 512
Db 362 YKIMRRIABEELTDASEIVITSTROEID-----EQWRLYDGFDPILERKLRARIKN 414
Oy 513 ---IDVPPKRNIVSPGADMEIYFPY-----TEKRRLKHFHPEIEDLYTKVENEH 562
Db 415 VSCYGRFMPMAVVIIPPGEFHHIIVPHGEDMDGETESSEDGKTPDPPIW-----AEI 465
Oy 563 LCVLNDNRKPILETWPRLDVNVNLTLGVWCGKNPKLRELANLIVVGGDRRK-ESKDLSE 621
Db 466 MRFENPRKPMILALARPDPKNLTLVKAFGECPRLRELANLTLIMGNRDNIDEMSTN 525
Oy 622 KAEMKMFELIDKYNLNGQFRWISSQMRIRNV-ELYRYICDTKGAFFVQPALYEAFGTLV 680
Db 526 SALLSLILKIMIDKYDLYGQVAV--PKHHKQSDVPDIYRLAAKTKGVFINPAFIEPFGTL 583
Oy 681 VEAMTCGLPTATCNGPAEIIYHKGSGFNIDPHYGDAQDILVDFEKKCKDPBHMCKI 740
Db 584 IEAAAYGIPWVATKXGPGVDIHRVLNDGNLVDPHDQAIADALL---KLVAHQQLMAKC 639
Oy 741 SQGLKRIEBEKYTW---KIYSERL 761
Db 640 RANGIAKNI-HLPSWPEHCKTYLSRI 663
```

Search completed: June 2, 2004, 14:53:49
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 20:48:25 ; Search time 1467 Seconds
(without alignments)
7601.581 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625

Sequence: 1 atgctcgcagcgtctctcac.....tcaaaaaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: N_Geneseq_29Jan04:*

1: _geneseq1980s:*\n2: geneseq1990s:*\n3: geneseq2000s:*\n4: geneseq2001as:*\n5: geneseq2001bs:*\n6: geneseq2002s:*\n7: geneseq2003as:*\n8: geneseq2003bs:*\n9: geneseq2003cs:*\n10: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624	100.0	2625	6	ABK52880 Cotton su
2	1499.4	57.1	2906	3	AAa67143 Eucalyptu
3	1497.8	57.1	2913	3	AAa67087 Eucalyptu
4	1497.8	57.1	3103	3	AAc62807 Sucrose s
5	1497.8	57.1	3103	6	ABK17072 Eucalyptu
6	1486.6	56.6	2427	7	ABa13815 Arabidops
7	1486.6	56.6	2427	7	ADa68333 Arabidops
8	1305.4	49.7	2714	9	ADc68423 Lolium pe
9	1305.4	49.7	2716	9	ADc68520 Lolium pe
10	1305.4	49.5	2829	6	ABK98516 CDNA sequ
11	1300.6	49.5	2494	3	AAc66090 Rice sucr
12	1296.6	49.4	2825	9	ADc68523 S. arundi
13	1295.4	49.3	2754	9	ADc68521 S. arundi
14	1287.6	49.1	2450	7	ADa71337 Rice gene
15	1287.6	49.1	2451	9	ADc08208 Rice DNA
16	1287.6	49.1	2451	9	ADc07855 Rice DNA
17	1287.6	49.1	2957	9	ADc68424 Lolium pe
18	1281	48.8	2908	6	ADd45852 Corn sucr
19	1274.8	48.6	2950	9	ADc68522 Lolium pe
20	1265.6	48.2	2360	3	AAc66096 PSS3 CDNA
21	1254	47.8	2746	6	ADd45851 Corn sucr
22	1124.2	42.8	2334	7	ADa71106 Rice gene
23	1124.2	42.8	2334	9	ADc07857 Rice DNA

24	1124.2	42.8	2334	9	ADc08296 Rice DNA
25	1115	42.5	2430	6	ABz14610 Arabidops
26	1104.8	42.1	2394	7	ABz76372 A. thalia
27	1104.6	42.1	2757	6	ADd45856 Corn sus3
28	1097	41.8	2418	7	ABz76371 A. thalia
29	1096.2	41.8	2737	6	ADd45849 Corn sucr
30	1053.4	40.1	2563	2	AAQ08005 Sucrose-s
31	1048.6	39.9	2412	7	ADa69581 Rice gene
32	1008.4	38.4	2210	7	ADa69597 Rice gene
33	740.2	28.2	2397	7	ADa70504 Rice gene
34	740.2	28.2	2397	9	ADc07859 Rice DNA
35	703.4	26.8	2361	9	ADc07861 Rice DNA
36	652.8	24.9	1374	9	ADc08717 Wheat DNA
37	527	20.1	1209	6	ABK98492 Consensus
38	491.6	18.7	2700	2	AAV21641 Arabidops
39	473.2	18.0	1145	3	AAa67119 Pinus rad
40	451	17.2	1169	3	AAa67120 Pinus rad
41	417	15.9	1064	9	ADc08723 Wheat DNA
42	334.4	12.7	2492	6	ADd42841 Maize suc
43	332.2	12.7	676	6	ABK98487 CDNA sequ
44	302.4	11.5	697	6	ABK98493 CDNA sequ
45	301.6	11.5	1546	9	ADc07853 Rice DNA

ALIGNMENTS

RESULT 1
ABK52880
ID ABK52880 standard; DNA; 2625 BP.
XX
AC ABK52880;
XX
DT 13-AUG-2002 (first entry)
XX
DE Cotton sucrose synthase Susy gene.
XX
KW Cotton; Sucrose synthase; Susy; ds; fibre; seed; transgenic; plant; gene.
XX
OS Gossypium hirsutum.
XX
FH Key Location/Qualifiers
FT CDS 1..2421 a
FT /*tag= /product= "Sucrose synthase protein"
XX
XX
XX MO200245485-A1.
XX
XX 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-AU001580.
XX
XX 08-DEC-2000; 2000AU-00001975.
XX
XX 08-DEC-2000; 2000US-0251852P.
XX
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
XX Llewellyn D, Furbank R, Ruan Y;
XX
XX WPI; 2002-463779/49.
XX
XX P-PSDB; AAU97898.
XX
XX
XX Altering fibre development or properties of a fibre producing plant by
XX modulating sucrose synthase activity and/or expression in such plants,
XX useful for enhancing fibre yield and quality and for increasing seed
XX size.
XX
XX Claim 5; Page 53-56; 62pp; English.
XX
XX This invention relates to a novel method for altering fibre development
XX or properties of a fibre producing plant by modulating sucrose synthase
XX (Susy) activity and/or expression in such plants. The invention also
XX comprises a fibre producing plant comprising a chimeric gene in its

CC genome, the seeds of the plant and fibre isolated from the plant. The
CC method is useful for altering fibre development or properties of a fibre
CC producing plant like cotton plant. Therefore, the method is useful for
CC enhancing fibre yield, enhancing fibre quality and for increasing seed
CC size in a fibre producing plant. The present sequence represents the DNA
CC encoding the cotton sucrose synthase (Susy) protein used to create the
CC transgenic plant of the invention

XX Sequence 2625 BP; 730 A; 560 C; 605 G; 729 T; 0 U; 1 Other;

Query Match 100.0%; Score 2624; DB 6; Length 2625;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCTGAGCGTGTCTCACTCGCGTCCACAGTCTCCGTGAGCGTGTGGATGAGACCTT 60
DB 1 ATGCTGAGCGTGTCTCACTCGCGTCCACAGTCTCCGTGAGCGTGTGGATGAGACCTT 60
QY 61 CTGCTCAAGAAACGAGATTTTGGCTTGTCTTCAAGATCGAAGGCAAAAGAA 120
DB 61 CTGCTCAAGAAACGAGATTTTGGCTTGTCTTCAAGATCGAAGGCAAAAGAA 120
QY 121 ATTCTGCAACCAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAGAGAAAG 180
DB 121 ATTCTGCAACCAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAGAGAAAG 180
QY 181 AAGCTCGTAATGAGTGCATTTTGAAGTATGAAGCTAGTACAGAAAGCATGCTTG 240
DB 181 AAGCTCGTAATGAGTGCATTTTGAAGTATGAAGCTAGTACAGAAAGCATGCTTG 240
QY 241 CTTCCATGAGTTCATCTTCTGTTCTCAAGGCGTGTGTTGGAGTACATTAAGTG 300
DB 241 CTTCCATGAGTTCATCTTCTGTTCTCAAGGCGTGTGTTGGAGTACATTAAGTG 300
QY 301 AATGTTACCGCTTGTGTTGAGAACTCACTGTTGAGTATCTTCAAGGAA 360
DB 301 AATGTTACCGCTTGTGTTGAGAACTCACTGTTGAGTATCTTCAAGGAA 360
QY 361 GAGCTTGTGAGTGAAGTTCAATGSAACCTTGTGTAATGAAATTTGAGCCCTTC 420
DB 361 GAGCTTGTGAGTGAAGTTCAATGSAACCTTGTGTAATGAAATTTGAGCCCTTC 420
QY 421 AACTCATATTCCTCCCGCCCAACTCTTCAAAATCATTGGTAATGAGTGAATCTTA 480
DB 421 AACTCATATTCCTCCCGCCCAACTCTTCAAAATCATTGGTAATGAGTGAATCTTA 480
QY 481 AATGTCACCTTTGGGCAAAATGTTCCATGACAAAGGAGACGACCCCTTGCTGAA 540
DB 481 AATGTCACCTTTGGGCAAAATGTTCCATGACAAAGGAGACGACCCCTTGCTGAA 540
QY 541 TTCTCTCAGAGTCAATGTCACAAAGGCAAGAACTGATCTTGGTACCTTACCTCT 600
DB 541 TTCTCTCAGAGTCAATGTCACAAAGGCAAGAACTGATCTTGGTACCTTACCTCT 600
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAAGAGATCTTGGAACTCTCTCT 660
DB 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAAGAGATCTTGGAACTCTCTCT 660
QY 661 GAGGACCAATGTCGAAATTTGAAACCGGTTCCAGAAATCGGTTGGAAAGAGTTGG 720
DB 661 GAGGACCAATGTCGAAATTTGAAACCGGTTCCAGAAATCGGTTGGAAAGAGTTGG 720
QY 721 GGTGACACCGGAGACGGGTCTCGAGATGATCAACTCTTTGGATCTTCTTGAAGCA 780
DB 721 GGTGACACCGGAGACGGGTCTCGAGATGATCAACTCTTTGGATCTTCTTGAAGCA 780
QY 781 ACTGATCTCTTGAACCTTGAAGATCTTGGAGAAATCCCAATGATGTAATGTTGTG 840
DB 781 ACTGATCTCTTGAACCTTGAAGATCTTGGAGAAATCCCAATGATGTAATGTTGTG 840
QY 841 ATTCTCACTCCCAAGGATTAATCTTCAAGAAATGTTTGGGGTATCCGGAACCGGT 900
DB 841 ATTCTCACTCCCAAGGATTAATCTTCAAGAAATGTTTGGGGTATCCGGAACCGGT 900
```

```
QY 901 GGGCAGGTTGTTACATCTTGTGATCAAGTCGAGCTTTGGAGAAATGATGCTCTCCGT 960
DB 901 GGGCAGGTTGTTACATCTTGTGATCAAGTCGAGCTTTGGAGAAATGATGCTCTCCGT 960
QY 961 ATAAAGAAACGAGTCAACATCACCCCTCGAATCCTCTATTTACTAATCTTCTCT 1020
DB 961 ATAAAGAAACGAGTCAACATCACCCCTCGAATCCTCTATTTACTAATCTTCTCT 1020
QY 1021 GATGCTGTGGAACAACATGCGGTCAACGACTTGAAGAAATATACGAAACAGACACTCG 1080
DB 1021 GATGCTGTGGAACAACATGCGGTCAACGACTTGAAGAAATATACGAAACAGACACTCG 1080
QY 1081 GATATCTTGAAGTACCTTCAAGAACGAAAGGAAATGTTGAAATATGATCTCAAGA 1140
DB 1081 GATATCTTGAAGTACCTTCAAGAACGAAAGGAAATGTTGAAATATGATCTCAAGA 1140
QY 1141 TTTGAAAAAATCTGCGCATCTTGGAAAACCTTACACAGAGATGTGTCTATGAAATCTCC 1200
DB 1141 TTTGAAAAAATCTGCGCATCTTGGAAAACCTTACACAGAGATGTGTCTATGAAATCTCC 1200
QY 1201 AAAGAGTTGACGAGCAAGCAGATCTGATCATCGGAAACNACAGCGACGCAATATCTGT 1260
DB 1201 AAAGAGTTGACGAGCAAGCAGATCTGATCATCGGAAACNACAGCGACGCAATATCTGT 1260
QY 1261 GCTCTCTTGTGCGCAATAAATTAAGTGTGACACAGTGCACATCGCCATGCTTTGGAG 1320
DB 1261 GCTCTCTTGTGCGCAATAAATTAAGTGTGACACAGTGCACATCGCCATGCTTTGGAG 1320
QY 1321 AAGACAAAATATCAATTCAGATATCTATTTGGAAGAAAGTTGAAACAAATACATTC 1380
DB 1321 AAGACAAAATATCAATTCAGATATCTATTTGGAAGAAAGTTGAAACAAATACATTC 1380
QY 1381 TCTTGCCAATTTACAGTGAATCTTTTTCGAATGAAACCATACAGATTCATCATCACAGT 1440
DB 1381 TCTTGCCAATTTACAGTGAATCTTTTTCGAATGAAACCATACAGATTCATCATCACAGT 1440
QY 1441 ACTTTCAGGAAATTCGAGAGACGAAACACATGTTGTCAATTCAGAGCCACATCTGCT 1500
DB 1441 ACTTTCAGGAAATTCGAGAGACGAAACACATGTTGTCAATTCAGAGCCACATCTGCT 1500
QY 1501 TTCACTTCTTCCGTGCTCTACCGGTGTTGACAGATGATGATGATGATGATGATGATG 1560
DB 1501 TTCACTTCTTCCGTGCTCTACCGGTGTTGACAGATGATGATGATGATGATGATGATG 1560
QY 1561 AACATGTTTCCCTGGTGTGATGAGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 AACATGTTTCCCTGGTGTGATGAGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 AGGTTGAAGATTTCCATCTGAGATGAAAGACCTTCTTTACACCAAGTTGAGATGAA 1680
DB 1621 AGGTTGAAGATTTCCATCTGAGATGAAAGACCTTCTTTACACCAAGTTGAGATGAA 1680
QY 1681 GAACACTTATGTTGTGCTCAATGACCGGACAAACCAATCTTGTTCACAAAGCCAGCTT 1740
DB 1681 GAACACTTATGTTGTGCTCAATGACCGGACAAACCAATCTTGTTCACAAAGCCAGCTT 1740
QY 1741 GATCGTGTCAAGAACTTAAACCGGACTGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGT 1800
DB 1741 GATCGTGTCAAGAACTTAAACCGGACTGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGT 1800
QY 1801 GAGTTGGCTTAACTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 1801 GAGTTGGCTTAACTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 GAGAAAGCTGAATGAAAGAAATGTTTGAAGCTATGACAAAGTACAACTTGAACCGGCA 1920
DB 1861 GAGAAAGCTGAATGAAAGAAATGTTTGAAGCTATGACAAAGTACAACTTGAACCGGCA 1920
QY 1921 TTCAATGATGATATCATCTCAATGAAACAGAAATCGAAATGTTGAACCTTACCGATACAT 1980
DB 1921 TTCAATGATGATATCATCTCAATGAAACAGAAATCGAAATGTTGAACCTTACCGATACAT 1980
```

Qy	1981	TGCGACAGAAAGGCGCTTTGTACAGCTGATGTATGAAGCCTTTGATTCAGCTT	2040
Db	1981	TGCGACAGAAAGGCGCTTTGTACAGCTGATGTATGAAGCCTTTGATTCAGCTT	2040
Qy	2041	GTGAGCGCAATGACTTGGCGTTTGGCCAACTTCGCAACCTGTAAACGCTGACAGCCGAG	2100
Db	2041	GTGAGCGCAATGACTTGGCGTTTGGCCAACTTCGCAACCTGTAAACGCTGACAGCCGAG	2100
Qy	2101	ATTATATGTCATGGGAAATCTGGTTTCAACATTCATCTTAACATGTGATCAGCTGCT	2160
Db	2101	ATTATATGTCATGGGAAATCTGGTTTCAACATTCATCTTAACATGTGATCAGCTGCT	2160
Qy	2161	GACATATCTCGTCGATTTCTTTGAAAGGTAAAGATTCATCTCACTGGGATTAAGTCT	2220
Db	2161	GACATATCTCGTCGATTTCTTTGAAAGGTAAAGATTCATCTCACTGGGATTAAGTCT	2220
Qy	2221	TCCCAAGAGAGCCTTGAAACGAATAGAGAGAGATATACATGAGAAATTATCTCGAGAGAG	2280
Db	2221	TCCCAAGAGAGCCTTGAAACGAATAGAGAGAGATATACATGAGAAATTATCTCGAGAGAG	2280
Qy	2281	CTATTGACCCCTGACAGAGAGTATGGAATTCGGAAGCATGTTTCCAACTTGAACGCCGT	2340
Db	2281	CTATTGACCCCTGACAGAGAGTATGGAATTCGGAAGCATGTTTCCAACTTGAACGCCGT	2340
Qy	2341	GAGAGTCTGCTGTACTCTGGAATGTTTATGCTCTTAAGTACCGTAAGCTGGCTGAATCA	2400
Db	2341	GAGAGTCTGCTGTACTCTGGAATGTTTATGCTCTTAAGTACCGTAAGCTGGCTGAATCA	2400
Qy	2401	GTTCCATTGGGACAGAGAGATTAATTGAACCTGTTAAATACATTTGGAGCCGGTTTCTTGG	2460
Db	2401	GTTCCATTGGGACAGAGAGATTAATTGAACCTGTTAAATACATTTGGAGCCGGTTTCTTGG	2460
Qy	2461	AGAAATTAATATCTGTTTGTATTAATTCATTGGAGAGCTCCTTGTATTTCAATCTGTCT	2520
Db	2461	AGAAATTAATATCTGTTTGTATTAATTCATTGGAGAGCTCCTTGTATTTCAATCTGTCT	2520
Qy	2521	TTTCTCTTTCTCTTTTTCGCGGCAATGTTGTAACATGGGGTGTGGCCCGCTCAATTC	2580
Db	2521	TTTCTCTTTCTCTTTTTCGCGGCAATGTTGTAACATGGGGTGTGGCCCGCTCAATTC	2580
Qy	2581	AGTTAAATATATGATGATCTTTGTTTTCAAAAAATTTTTTTTTTTTTTTTTTTTTT	2625
Db	2581	AGTTAAATATATGATGATCTTTGTTTTCAAAAAATTTTTTTTTTTTTTTTTTTTTT	2625
RESULT 2			
ID	AAA67143	standard; DNA; 2906 BP.	
AC	AAA67143;		
XX			
DT	31-OCT-2000	(first entry)	
DE	Eucalyptus grandis sucrose synthase nucleotide sequence SEQ ID NO:139.		
KW	Eucalyptus grandis; pinus radiata; Monterey pine; modification;		
KV	plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;		
XX	transgenic plant; de.		
OS	Eucalyptus grandis.		
XX			
PN	MO200022092-A2.		
XX			
PD	20-APR-2000.		
XX			
PF	08-OCT-1999; 99WO-NZ000169.		
XX			
PR	13-OCT-1998; 98US-00170862.		
PR	11-AUG-1999; 99US-0148426P.		
PA	(GENE-) GENESIS RES & DEV CORP LTD.		
PA	(FLEET-) FLETCHER CHALLENGE FORESTS LTD.		
XX			

Query	Match	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
Q1	1	ATGCGTGAAGCGTCTCTACCTCGCTCCACAGTCTCCGTAGACGTTTGATGAGACCTT	57.1%	1499.4	2906	1852	0	567	3	1
D1	233	ATGCGTGAATCCATGCTGACCTCGAAGCCACAGCTTCCGACGCTTTGACAGACCCCTC	76.5%	0						
Q2	61	CTTGCTCAGAGAAAGAGATTTTGGCCTTGCTCTGAGGATCGAAGGCAAGAAAGGA			120					
D2	293	TCTGCTACCGGCAACGATATTTGTGCGCTTCTTTAAGGTTGAAGCCAGGCAAGGCG			352					
Q3	121	ATTCGTCAACACATCAATATTTCTAGAGTTTGAAGTATCTCCCTGAAGAGACAGAAAG			180					
D3	353	ATCTTGACGGCCACACAGATTTTGTGATTTGAGGCACTCTGAGAGAGACAGAGA			412					
Q4	181	AAGCTCGTAAATGTCGATTTTGAAGTATTTGAAGGCTAGTCAGAAAGCATCTGTTG			240					
D4	413	AAGCTTCTGATGAGGCGCTTTGGAGTGTCTCAAAATCACATCAAGAAAGCGATGTGTG			472					
Q5	241	CTCTCAGAGGTGACCTTGTGCTGCTCCAGAGCGCTGTGTTTGGAGATCTAGAGTG			300					
D5	473	CTCTCAGAGGTGCTCTGCTGTGTCTCCAGAGCGGCGGTGTGGAGACATCCGTGTG			532					
Q6	301	AATGTCACGCGCTTGTGTGTGAGGAATCACTGTGTGATATCTCCATCTCAAGAA			360					
D6	533	AAGTCATAGGCGCTTGTCTTGAGGAATTTGAGAGTGTGTGATATCTGACATCTCAAGAA			592					
Q7	361	GAGCTTGTGAGGAAGTCAAAATGAAACTTTGTTTGAATTGAAATTTGAGCCCTTC			420					
D7	593	GAGCTTGTGAGGAAGTCAAAATGAAACTTTGTTTGAATTGAAATTTGAGCCATTC			652					
Q8	421	AACATCATATTCCTCCCGCCCAACTCTTCAAAATTCATTTGTAATAGTGTGAGATTCCTA			480					
D8	653	ACTGCTCTTTTTCGCGCCGCACTTTTCAAGTCTATTTGCAATAGGCGTCAAGTTTCTC			712					
Q9	481	AATGTCACCTTTTCGCAAAATTTGTTCCATGACAGAGACATCACTCTTGTCTGAA			540					
D9	713	AATGTCATCTCTCCGCTTAACCTCTTCAATGACAGAAAGCTTGACCTCTGTGTA			772					
Q10	541	TTTCCTCAGAGTCAATGTCAACAGGCGCAAGACATGATTTGAAATGACAGAAATCAAAAC			600					
D10	773	TTTCCTCAGAGTCAATGTCAACAGGCGCAAGACATGATTTGAAATGACAGAAATCAAAAC			832					
Q11	601	TTGAAATGCTTTTCAACATGTTTGAAGAAACACAGAGAGTACTTGTGATACCTACTCTCT			660					
D11	833	GTGTTCTCTCTCAACATGCTCTGAGGAAGGCGGAGAGTATCTGACCTGCTCAAAACC			892					

QY 661 GAGACACCATGTCGGAATTCGAAACCGGTTCCAGAAATCGGTTTGGAAAAGGTTGG 720
 DB 893 GAGACCCCGTACTCCAGTTCCAGACACAAATTCCAGAGATCGGAGCTCGAGCGGGGGTGG 952
 QY 721 GGTACACCGGAAACCGGTCTCGAAGTGAATCCATCTCTTTGGATCTTTGAAGCA 780
 DB 953 GGTACACCGGTGAGCGGCTCTCGAAGTGAATCCAGCTCTGTTGGATCTCTTGAAGCT 1012
 QY 781 ACTGATCTCTGACCTTGAAGAAATTCCTGGAGAAATCCCATGATGTTCAATGTTGTG 840
 DB 1013 CCGACCCGTTGACCTTCGAGAAATTTCTGGATAGGTTCCCATGATCTTCAAGTGTG 1072
 QY 841 ATTCTCATCTCCCAAGGATCTTGGCTCAAGACAAATGTTTGGGGTATCCGACACCGGT 900
 DB 1073 ATCATGTTCTCCCAAGGATCTTGGCTCAAGACACGCTTGGATATCCGATATCCGGT 1132
 QY 901 GGGCAGGTTGTTTACATTTGGATCAAGTCCGAGCTTTGGAGAAATGAGATCTCTCGT 960
 DB 1133 GGGCAGGTTGTTTACATTTGGATCAAGTCTGATCGCTTGAAGAAATGATCTTCAACGCT 1192
 QY 961 ATTAAGCAACAGGATCTCAATCACCCTCGAAATCTCATTTATTACTAGACTTCTTCT 1020
 DB 1193 ATTAAGCAACAGGATCTGATATTTATCTCTCGAATCTCATTTATCACTCGGCTTCTTCA 1252
 QY 1021 GATGCTGTGGAAACAAATGCGGTCAACGACTTGAAGAAATGATAGCAAGACACTCG 1080
 DB 1253 GACGCGGTTGAAACAACTGTGGCCAGCGCTTGAAGAAATGTTTGGGACCGAGTACTCC 1312
 QY 1081 GATATTTCTCGAAGTACCTTTGAAACAGAAAGGAAATGTTCCGAAATGATCTTCAAGA 1140
 DB 1313 CACATTTCTTGGCGCTCCCTTCAAGAAATGAGAGGAGTGTCCGCAAGTGAATTTCCGG 1372
 QY 1141 TTTGAAAAGTCTGGCCATTACTTGGAAACCTACACAGAGATGTTGCTCATGAATCTCC 1200
 DB 1373 TTGCG---AGTGTGGCCCTTATTTGGAAAGATACATGAGATGTGCGAGGAACTTGTCT 1429
 QY 1201 AAAGAGTTGACGCGACCGCAGATCTGATCATTCGGAACCAACAGCGACGGAATATGCTC 1260
 DB 1430 GGAAGGTTGACGCGGAGCCTGATCTGATCATTCGGAACCTACAGATGAGAAACATTTGTT 1489
 QY 1261 GCTTCCTTGTCTGCAATTAATTAAGTGTCAACAGTGCACATTCGCCCATGCTTTGGAG 1320
 DB 1490 GCTTCCTTGTGTAGCATTAATTAAGTGTCAACAGTGTCAACATTAAGCCCATGCTTCAG 1549
 QY 1321 AAAGCAAAATATCCGATTCAGATATCTATTGAAAGAACTTGAAGCAAAATACATTTTC 1380
 DB 1550 AAAGCAAAATATCCGATTCAGATATCTATTGAAAGAAATTTGAGAAAGAAATACATTTTC 1609
 QY 1381 TCTTGCCAAATTAAGCTGATCTTTTTCGAAATGAAACATACAGATTTGATCATCAACAGT 1440
 DB 1610 TCTTGCCAGTTTCACTGCTGATCTCATGCAATGAAACCAACGACTTATATCAACAGC 1669
 QY 1441 ACTTTCGAGAAATTTGCAAGACGACACTGTTGTCAATACGAGAGCCACACTGTCT 1500
 DB 1670 ACCTTTCGAGAAATTTGCAAGACGACACTGTTGTCAATACGAGAGCCACACTGTCAAC 1729
 QY 1501 TTCACTCTCTGAGTCTACCGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1730 TTCACTCTCTGAGTCTACCGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1789
 QY 1561 AACATTTGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 DB 1790 AACATTTGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1849
 QY 1621 AGGTTGAAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1850 CGGTTGAAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1909
 QY 1681 GAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 DB 1910 GAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1969

QY 1741 GATGCTGTCAAGAACTTAACCGGATCTGTGAGATGATGATGATGATGATGATGATGATGAT 1800
 DB 1970 GACCTGTGTCAAGAACTTGAACAGGCTTGTGAGTGTATGTGCAAGAACTCAAGTTGAGG 2029
 QY 1801 GAGTGTGTAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 DB 2030 GAACTGTCAATTTGT 2089
 QY 1861 GAGAGGCTGAAATGAAGAAATGTTTGAAGCTGATGATGATGATGATGATGATGATGATGAT 1920
 DB 2090 GAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2149
 QY 1921 TTCAAGATGATATCTCAATGAAACAGAAATCCGAAATGTTGAACCTTACCGATTAATT 1980
 DB 2150 TTCAAGATGATATCTCTCCAGATGAAACCGGATGATGATGATGATGATGATGATGATGAT 2209
 QY 1981 TGCACACGAAAGGTGCTTGTGATACAGCTGTGATGATGATGATGATGATGATGATGATGAT 2040
 DB 2210 TGTACACGAAAGGAGTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2269
 QY 2041 GTGAGGCAATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 DB 2270 GTTGAAGCAATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2329
 QY 2101 ATTAATGTGTCAAGGAAATCTGTTTCAACATTTGATCTTACATGATGATGATGATGATGAT 2160
 DB 2330 ATCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2389
 QY 2161 GACATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 DB 2390 GAGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2449
 QY 2221 TCCCAAGAGGCTTGAACCAATGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 2280
 DB 2450 TCAAGGCTGTCAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2509
 QY 2281 CTATTTGACCTGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 DB 2510 CTGTTGAACCTGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2569
 QY 2341 GAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 DB 2570 GAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2629
 QY 2401 GTTCATTTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2422
 DB 2630 GTTCCTCGGCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2651
 RESULT 3
 AAA67087
 ID AAA67087 standard; DNA; 2913 BP.
 XX
 AC AAA67087;
 XX
 XX
 DT 31-Oct-2000 (first entry)
 XX
 DE Eucalyptus grandis sucrose synthase nucleotide sequence SEQ ID NO:15.
 XX
 KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
 KW plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
 KW transgenic plant; ds.
 XX
 OS Eucalyptus grandis.
 XX
 EN W0200022092-A2.
 XX
 XX 20-Apr-2000.
 PD
 XX 08-Oct-1999; 99MO-NZ000169.
 PF
 XX 13-Oct-1998; 98US-00170862.
 PR 11-Aug-1999; 99US-0148426P.

XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR WPI; 2000-339328/29.
DR P-PSDB; AAB16282.
XX
PT New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant.
XX
PS Claim 1; Page 43-44; 301pp; English.
XX
CC The present invention describes isolated polynucleotides (PN) comprising
CC a sequence selected from one of 835 nucleotide sequences given in
CC AAB67073 to AAB67907, their (reverse) complements, sequences producing an
CC expectation (E) value of 0.01 or less compared to the 835 sequences,
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic to
CC the 835 sequences. The polynucleotides are used to modify the activity of
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
CC plant. They are especially used to modulate or alter the polysaccharide
CC content, composition or structure of the plant. AAB16268 to AAB16340 are
CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention
XX
SQ Sequence 2913 BP; 685 A; 709 C; 732 G; 787 T; 0 U; 0 Other;
Query Match 57.1%; Score 1497.8; DB 3; Length 2913;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
QY 1 ATGCTGAGCGTCTCTCACTCGCGTCCAGTCTCCGTGAGCGTTGGATGAGACCTT 60
DB 242 ATGCTGATCCAGATGTTGATCTGAGCAAGCCAGCTTCGCGAGCGTTTGAAGACCTCTC 301
QY 61 CTGCTCAAGAGAGAGATTTGGCGCTTCTCAAGGATCGAGCGGAAGAAAGGA 120
DB 302 TCTCTCAAGAGAGATTTGGCGCTTCTCAAGGATCGAGCGGAAGAAAGGA 361
QY 121 ATTCTGCAACACATCAATATTATCTAGATTGAGTTGAGTATCCCTGAGAGAAAGGA 180
DB 362 ATCTTGAGCGCCACACAGATTTTCTGAGATTGAGGCACTCTGAGAGAGAGAGCA 421
QY 181 AAGCTCGTAATGATGATTTTGAAGATTTGAAGGCTAGTCAAGAGAGAGATGCTGTG 240
DB 422 AAGCTTCTGATGAGGCGCTTGTGTAAGTCTCAAAATCCACTCAAGAGAGATGCTGTG 481
QY 241 CTTGCAAGGATTTGCACTTGTCTGTCTCAAGGCGTGTGTTGGAGATATATAGATG 300
DB 482 CTTCAAGGATTTGCTTGTCTGTCTGTCTCAAGGCGGCGTGTGGAACACATCCGTGTG 541
QY 301 AATGTTCAAGCGCTTGTGTGAGGAATCACTGTTGTGAGTATCTCCACTTCAAGGA 360
DB 542 AAGCTCAAGGCGCTTGTGTGAGGAATTTGAGAGTTGCTGAGTATCTCACTTCAAGGA 601
QY 361 GAGCTTGTGATGAGATTTGAATGAAATCTTGTGTTGATGATTTTGAAGCTTTC 420
DB 602 GAGCTTGTGATGAGATTTGAATGAAATCTTGTGTTGATGATTTTGAAGCTTTC 661
QY 421 AAGCTCAATATCCCGCGGCAATCTTCAAAATCCATTTGTAATGATTTTGAAGCTTTC 480
DB 662 AAGCTCAATATCCCGCGGCAATCTTCAAAATCCATTTGTAATGATTTTGAAGCTTTC 721
QY 481 AATGTTCACTTTGCGCAAAATTTGTTCAATGACAAGAGAGATGACCTTTGCTGAA 540
DB 722 AATGCGCATCTCTCGCTAGCTCTTCAATGACAAGAGAGATGACCTTTGCTGAA 781
QY 541 TTCTCAAGATGCTATGTCACAAAGGCAAGATGATGTTGAATGACAGATTCAGAAC 600
DB 782 TTCTCAAGATGCTATGTCACAAAGGCAAGATGATGTTGAATGACAGATTCAGAAC 841

QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTATACCTTCT 660
DB 842 GTGTTCTCCCTCCCAACATGTTCTGAGAGAGAGAGATCTGACCTGCTCAAAACC 901
QY 661 GAGACACCATGTCGAAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGG 720
DB 902 GAGACACCGGTTCTCCAGTTGAGACACAAGTTCCAGAGATCGGAGCTCGAGCGGGGTGG 961
QY 721 GGTACACCGGAGAACCGGTGCTGAGATGATGATTCATCTCTTTTGGATCTTGAAGCA 780
DB 962 GGTACACCGGAGAACCGGTGCTGAGATGATGATTCATCTCTTTTGGATCTTGAAGCA 1021
QY 781 ACTGATCTTTCACACCTTGAAGATCTTGTGAGAAATCCCATGAGTTCATATGTTTG 840
DB 1022 CCGACCGGTCGATCTCTGAGAAATTTTGGATGAGTTCCCATGAGTTCATATGTTTG 1081
QY 841 ATTCTCACTCCCAAGAGATCTTGTGCTCCAGACATGTTTGGGATATCCGACACCGGT 900
DB 1082 ATCATGTTCTCCCAAGAGATCTTGTGCTCCAGACATGTTTGGGATATCCGATACCGGT 1141
QY 901 GGGCAGGTTGTTTACATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1142 GGGCAGGTTGTTTACATCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1201
QY 961 ATTAAGCAAGAGATCTCAATCATGACCCCTGAAATCTCATTTACTAGACTTCTTCT 1020
DB 1202 ATTAAGCAAGAGATCTCAATCATGACCCCTGAAATCTCATTTACTAGACTTCTTCT 1261
QY 1021 GATGCTGCGAACAACATGCGGTCAACGATGAGAAATGATGAGAAAGAGACATG 1080
DB 1262 GACCGGTTGGAACAACATGCGGTCAACGATGAGAAATGATGAGAAATGATGAGAAATG 1321
QY 1081 GATATTTCTGAGATCTTCAAGACAGAAAGGAAATGTTTCCAAATGATCTCAAGA 1140
DB 1322 CACATTTCTGAGATCTTCAAGACAGAAAGGAAATGTTTCCAAATGATCTCAAGA 1381
QY 1141 TTTGAAAGATCTGAGATCTTCAAGACAGAAAGGAAATGTTTCCAAATGATCTCAAGA 1200
DB 1382 TTTG--AGGTTGAGCTTATTTGAAAGATCTGAGAAATGATCTCAAGATGATGATGAT 1438
QY 1201 AAGAGTTGACAGGACAGCGCATCTGATCATCGAAACCAACAGCGAGCAATATCTG 1260
DB 1439 GAGAGTTGACAGGACAGCGCATCTGATCATCGAAACCAACAGCGAGCAATATCTG 1498
QY 1261 GCTTCCTTCTGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1499 GCTTCCTTCTGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
QY 1321 AAGCAAAATATCAAGATCTGATATCAATTTGAAGAAAGCTTGAAGCAAAATCAATTC 1380
DB 1559 AAGCAAAATATCAAGATCTGATATCAATTTGAAGAAAGCTTGAAGCAAAATCAATTC 1618
QY 1381 TCTTCCAAATTAAGATCTTATTTGCAATGAAACCAATATCATCATCATCATCATCAT 1440
DB 1619 TCTTCCAAATTAAGATCTTATGATCATGATCATGATCATGATCATGATCATGATCATG 1678
QY 1441 ACTTTCAGAAATTTGCAAG 1500
DB 1679 ACTTTCAGAAATTTGCAAG 1738
QY 1501 TTCACTCTCTGATCTCTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1739 TTCACTCTCTGATCTCTACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1798
QY 1561 AACATGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1799 AACATGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1858
QY 1621 AGGTTGAAGATTTTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 1859 CGGTTGAAGATTTTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1918

QY	1681	GAACACTTATGTGCTCAATGACCGCAACAAGCAATCTGTTCACATGCCAAGCTT	1740
Dp	1919	GAACACTTGTGTGTTGAAAGATPAAAGAAAGCCTATATTTCCACCATGGCAAGCTG	1978
QY	1741	GATCGTGTCAAGAACTTAAACCGGACTCGTGCAGTGGTGCAGCAAGAAACCAAGTTGCGT	1800
Dp	1979	GACCGTGTCAAGAACTTGAACAGGGCTGTTGTGATGGTATGGCAAGAACTCCAAGTTGAGG	2038
QY	1801	GAGTTGGCTAAACCTCGTAGTTGTAGTGTGTATAGCGAAAGGAATCTAAAGATTTTGGAA	1860
Dp	2039	GAACCTCGCAACTTGTGTCGTGGTTGGAAGTACACGGAAGGAAGTATCGAAGCACTTGGAA	2098
QY	1861	GAGAAGGCGTGAATGAAAGAAATGTTTGAAGCTGATCGACAAGTACAACTTGAACGGCCAA	1920
Dp	2099	GAGCACTGTGATGATGAAGAAATATGACACTTCATCGAAAGTACACAGCTBAATGGCCAG	2158
QY	1921	TTCCAGATGATATCATCTCAATATGACACGAATCCGAATGTTGAACCTTACCGATACAT	1980
Dp	2159	TTTCAGGTGGATTTCTCTCCAGATGAAACGGGGTGAAGAAATGGAAGCTCTACCGGTACATC	2218
QY	1981	TGCGACACGAAAGGTGCTTTGTATACAGCTGTGCATTTGATGAAGCTTTGCAATTGACGTT	2040
Dp	2219	TGTACACAGAAAGGAGATCTTGTTCAACCGGCTATCTATGAAGCTTTGCGGTTGACCGTG	2278
QY	2041	GTGGAGGCAATGACTGTCGGTTTGCCAACTTCGCCAACCCTGAACGGTGAACGACGCGAG	2100
Dp	2279	GTTAAGGCGCATGACTTGTGATGTGCCAACTTTGTCCACTTGCAAATGGTGAACCAAGCTGAG	2338
QY	2101	ATTATGTGCATGGGAAATCTGTTTTCACATTTGATCCTTACCAATGGTATCAAGCTGCT	2160
Dp	2339	ATCATTTGTGCATGGCAAAATCGGGCTACACATTAATCCTTAACATGTGTGAACCAAGCGGCC	2398
QY	2161	GACATPACTCGTCGATTTCTTTGAAAAGTGTAAAGAAAGATCCATCTCACTGGGATTAAGATC	2220
Dp	2399	GAGCTTCTTTGATGACTTCTTCAACAAGTGCAAAGATTGCAACAGTCCCACTGGGAGCAAGATC	2458
QY	2221	TCCCAAGAAGCCTGAAACGAATAGAGAAGAAAGTATACATGGAAGATTTAATCGGAGAGA	2280
Dp	2459	TCAAAAGGCTGCATGACAGAAATTTGAAGAAAGATTAACATGAAAAATATATTTCTGAAGG	2518
QY	2281	CTATTTGACCCCTGACAGAGAGTATATGAAATCTTGAAGACATGTTTCCAACTTGAACGCGGT	2340
Dp	2519	CTGTTTGAACCTGACCTGCCGTGTATATGGCTTCTGGAAGCATGTGACATACTTGAATCGGGCC	2578
QY	2341	GAGAGTCGTGCTTAACCTTGAGATGTTTATGCTCTTAAATGACGTAAGCTGGCTGAATCA	2400
Dp	2579	GAGAGTCGCGCGTACCTTGAATATGTTCTATGCTCCCTCAAGTATCGCCACATGGCAAGCTCT	2638
QY	2401	GTTCCATTTGACAGAGAGTAAA 2422	
Dp	2639	GTTCCTCCGGCTGTGAGGTAAA 2660	

	RESULT 4
AAAC62807	
ID	AAAC62807 standard; DNA; 3103 BP.
XX	
AC	AAAC62807;
XX	
DT	02-FEB-2001 (first entry)
XX	
DE	Sucrose synthase promoter coding sequence #1.
XX	
KM	Promoter; eucalyptus; pine; gene transcription; ds
XX	
OS	Eucalyptus grandis.
XX	
PN	W0200058474-A1.
XX	
PD	05-OCT-2000.
XX	
PF	24-FEB-2000; 2000W0-NZ000018.
XX	

	XX	PR	25-MAR-1999;	99US-00276599.	
	XX	PR	30-JUN-1999;	99US-0146591P.	
	XX	PA	(GENE-) GENESIS RES & DEV CORP LTD.		
	XX	PA	(FLET-) FLETCHER CHALLENGE FORESTS LTD.		
	XX	P1	Perera R, Rice SJ, Eagleton CK;		
	XX	WPI	2000-647236/62.		
	DR	DR	P-PsDB; AAB28141.		
	XX	Novel promoter sequences useful for modulating transcription of plant DNA			
	PT	sequences of interest and production of polypeptides.			
	XX	Claim 1, Page 61; 93pp; English.			
	CC	The present invention relates to promoter sequences from eucalyptus and			
	CC	pine. The present sequence is one such promoter. This sequence is useful			
	CC	for modulating the transcription of DNA sequences of interest. The			
	CC	sequences may also be used to tag or identify an organism or its			
	CC	reproductive material			
SQ	Sequence 3103 BP; 756 A; 740 C; 772 G; 835 T; 0 U; 0 Other;				
Query Match	57.1%; Score 1497.8; DB 3; Length 3103;				
Best Local Similarity	76.4%; Pred. No. 0;				
Matches 1851; Conservative	0; Mismatches 568; Indels 3; Gaps 1;				
QY	1 ATGGCTAGAGGTGCTCACTCGCGTCCACAGTCTCCGTGAGCGTTTGATGAGACCCTT	60			
Db	447 ATGGCTATATGCATTGTGACTCAGAAGCACAGCTTCGCCAGCGTTTGAGAGACCCCTC	506			
QY	61 CTTCGTACAAGAAAGAGATTTTGGCGTTCCTCAAGATCGAGGGCAAAGAAAAGAA	120			
Db	507 TCCTCATCGCAACGATATTGTGGCGCTTCCTTCAAGGGTTGAAGCCAAGGCAAAAGGC	566			
QY	121 ATTCTGCAACACCATCAATTATTTCTAGAGTTTGAAACTATCCCTGAAAGAAACAAAG	180			
Db	567 ATCTTGGAGGCCAACAGATTTTTTGTGAGTTTGAAGGCCAATCTCTGAGGAAGACAGCA	626			
QY	181 AAGCTGCTAATGGTGTGATTTTTTGAAGTAATGAAGGCTAGTCAAGAACGATCGTGTG	240			
Db	627 AAGCTTTTATATGGGCGCTTTGGTAGAGTCTCTCAATCTCACTCAAGAAAGCATTTGTGCG	686			
QY	241 CCTGCATGGGTGCACTTGCTGTTCGTCAAGGCGTGGTGGTGGAGTACATTAGAGTGTG	300			
Db	687 CCTCATGGGTGTGCTTGTGCTGTTCGTCAAGGCGGCGGTGGAGACATCGGTGTG	746			
QY	301 AATGTTCAGCGCCTTGTGTGTGAGAACTCACTGTGTGATATCTTCACTTCAAGAA	360			
Db	747 AACCTCATATCGCTTGTGTCTTGAAGCAATTGAGAGTTCTGTGATCTTCAAGAA	806			
QY	361 GAGCTTGTATGGAATGGAATGAAACTTTGTTTGAATGGAATTTGAGCGCTTC	420			
Db	807 GAGCTTGTATGGAATGGAATGGAATGGAATCTTTGTGCTTGAAGCTTGTGAGCCATTTC	866			
QY	421 AACTCATCATTTCCCCCGCCCAACTCTTCAAAATCAATTGGTAAATGATGAGATTCTTA	480			
Db	867 ACTGCGCTTTTCCGGCGCCGACCTTTTCCAAGCTAATTTGGCAATGGGTGAGATTCTC	926			
QY	481 AATGTCACCTTTTGGCAAAATTTGTTCCATGACAGAAGACATGACCTTTGCTGAA	540			
Db	927 AATGTCACATCTTCGCTAAGCTCTTTCATGACAGAAAGAAAGCTTGCACCCCTCTCTGAA	986			
QY	541 TTCTTCAGATTCATTTGTCACAAGGGAACAACATGATTTGATGTGACAGAAATTCAGAAC	600			
Db	987 TTCTTCAGATTCATTTGTCACAAGGGAACAACATGATTTGATGTGACAGAAATTCAGAA	1046			
QY	601 TTGAATCTCTTCAACATGTTTGAAGAAAGACAGAGATATCTTTGATCCCTACTCTCT	660			
Db	1047 GTGTTCCTCCTCCAACTGTCTCTGAGAAAGCGGAGAGATATCTGACTCGCTCAAAACC	1106			
QY	661 GAGCACCATATGCGCAATTCGAACACCGGTTCAGAGAAATCGGTTTGAAAAGAGTTTG	720			

[illegible]

Dd		2184	GACCGTGTCAAGAACTTGACAAGGCGCTTTGGTAGTGATATGCCAAGAATCCAAAGTTGAGC	2243
Oy		1801	GAGTTGGCTAACCTTCGTAGTTGGTAGTGATATAGCGAAAAGAACTTAAGAATTGGAA	1860
Dd		2244	GAACCTCCGCAACTTGCTGTGTGTGGAGGTGACAGAGGAAAGAAATTCGAAAGCATTTGGAA	2303
Oy		1861	GAGAAAGCTGAAATGAAGAAAAATGTTGAAGTATCCAGATGAACCTTGAAGGCCCA	1922
Dd		2304	GAGCAGCTTGAGATGAAGAAAAATGTACGACTTCATCGAAAAGTACAGCTGAATGGCCAG	2363
Oy		1921	TTCCAGATGATATATCTCAATATGAACGAATCCGAAATGTGAACTTTACCATACTT	1980
Dd		2364	TTCCAGTGTGATTTCTCCCNATATGACCGGATGAGAAATGGAAGACTTACCGCTACATC	2423
Oy		1981	TGCACACGAAAGGTGCTTTGTACAGCTGTCATTTATGAAGGCTTTGANTGACAGTT	2040
Dd		2424	TGTGACACGAAGGAGATCTTGCTTCAACCGGCTATATATGAAGCTTTGSGTTGACCGTG	2483
Oy		2041	GTGNAAGCAATGATTTGGCGTTTTGCCCAATTCGCCAACCTGTACGGTGAACGACCGAG	2100
Dd		2484	GTTGAGGTCATGACTTGTGGAATGTCGCAACTTTGCCACTTGCAATGGTGAACGACTGAG	2543
Oy		2101	ATTATTTGTCATGAGGAATCTGTTTTCAACATTGATCCTTACATGGTATCAAGCTGCT	2160
Dd		2544	ATCATTTGTGATGGCAAAATCGGGCTACACATTTATCTTTACATGTTGACCAAGCGCC	2603
Oy		2221	TCCCAAGAGGCTTGAACCAATGAGAGAGAAATATCATATGAAGATTACTCGGAGGA	2280
Dd		2664	TCAAAGGGTGCATGACAGAGAAATGAAGAAAGATATCATGGAATAATATTTCTGAGAGG	2723
Oy		2281	CTATTGAACCCGACAGAGGTATGATTTCTGGAAGCATGTTTCCAACCTTGAAGCGCGT	2340
Dd		2724	CTGTTGAACCTGACTGCCGTGTATGCTTCTTGGAAGCATGTGACTTAACCTTGAATCGGCGC	2783
Oy		2341	GAGAGTGTGCTTACCTVTGAGATGTTTATGCTCTTAAGTACCCTGAAGCTGGCTGAATCA	2400
Dd		2784	GAGAGTGGCGGTACTTGAATATTTCTATGSCCTCAAGTATGCGCCACCTGGCACAGTCT	2843
Oy		2401	GTTCCATTGSCAGAGAGTAAA	2422
Dd		2844	GTTCTCCGGCTGTGCGAGTAAA	2865
RESULT 5				
ABK17072				
ID	ABK17072	standard; cDNA; 3103 BP.		
XX	ABK17072;			
AC				
XX				
DT	26-MAR-2002	(first entry)		
XX				
Dd	Bucalyptus grandis promoter polynucleotide #28.			
XX				
KW	Promoter; pine; leaf; root; flower; pollen; bud; meristem; xylogenesis;			
KM	temporally regulated promoter; Pinus radiata; Bucalyptus grandis; ss;			
KM	PCR primer.			
OS	Bucalyptus grandis.			
XX				
PN	MO2001.98485-A1.			
PD				
XX	27-DEC-2001.			
PF				
XX	20-JUN-2001; 2001WO-NZ000115.			
PR	20-JUN-2000; 2000US-00598401.			
PR	28-NOV-2000; 2000US-00724624.			
XX				
PA	(GENE-) GENESIS RES & DEV CORP LTD			

PA (FLET-) FLETCHER CHALLENGE FORESTS IND LTD.

XX Perera R, Rice S, Eagleton C, Lasham A;

XX WPI: 2002-114583/15.

DR P-PSDB; AAU80759.

XX Novel polynucleotide promoter sequences from pine and Eucalyptus useful
PT for modifying expression of endogenous and/or heterologous
PT polynucleotides in transgenic plants.

XX Claim 1; Page 76-77; 121pp; English.

XX The invention relates to isolated promoter sequences from *Pinus radiata*
CC and *Eucalyptus grandis*, comprising a sequence chosen from leaf-, root-,
CC flower-, pollen-, bud-, meristem-specific promoters or temporally
CC regulated promoters such as xylem-specific promoters. The promoter
CC polypeptides and their related polynucleotides are useful in the
CC production of genetic constructs, used for modifying gene expression in a
CC target organism, in particular a plant. The method is useful for
CC modifying expression of a polynucleotide that comprises an intron
CC sequence, through removal of the intron sequence. The method is useful
CC for modifying growth and development of plants, and cellular responses to
CC external stimulus, such as environmental factors and disease pathogens.
CC The sequences are useful in genome and physical mapping, in positional
CC cloning of genes, in various assays to determine biological activity, to
CC raise antibodies, to isolate corresponding interacting proteins and other
CC compounds, and to quantitatively determine levels of interacting proteins
CC and *Eucalyptus grandis* polynucleotides and PCR primers used in the method
CC of the invention

XX Sequence 3103 BP; 756 A; 740 C; 772 G; 835 T; 0 U; 0 Other;

Query Match 57.1%; Score 1497.8; DB 6; Length 3103;

Best Local Similarity 76.4%; Pred. No. 0; Mismatches 568; Indels 3; Gaps 1;

Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;

QY 1 ATGGCTGAGCGTGTCTCTCACTCGCTCCAGTCTCCGTGAGCGTTTGATGAGACCTT 60
DB 447 ATGGCTGATGCGATGTTGATCTGAGACCAAGCGCTTGGAGAGAGACCTTC 506
QY 61 CTGTGCTCAGAGAGAGATTTGGCTTGTCTCTGAGAGATCGAGGCGAAGAAAGGA 120
DB 507 TCTGTCTCAGCGCAAGATTTGGCTTGTCTCTGAGAGATCGAGGCGAAGAAAGGC 566
QY 121 ATTGCGAACAACATCAATATTATTCTAGATTTGAGATCCCGAAGAGAACAGAAAG 180
DB 567 ATCTTGACGCGCCACCAATTTTGTCTAGATTTGAGAGATCTTGAGAGAGAGAGACA 626
QY 181 AAGCTCGCTAAATGTCATTTTGTGAAATGTAAGGCTAGTCAAGAGACGATGCTT 240
DB 627 AAGCTTCTTAATGGGGCTTTGGTGAATGTCCTCAATCACTCAAGAGAGCAATGTCG 686
QY 241 CCTCATAGGGTTGCACTTGTCTGTTCAGAGGCTGTGTTGGAGATCAATTAGAGTG 300
DB 687 CCTCATAGGGTTGCTTGTCTGTTCAGAGGCTGTGTTGGAGATCAATTAGAGTG 746
QY 301 AATGTTCAAGCGCTTGTGTTGAGAGATCACTGTTGCTGAGATCTCACTTCAAGGAA 360
DB 747 AAGCTCAATGCGCTTGTGTTGAGAGATCACTGTTGCTGAGATCTCACTTCAAGGAA 806
QY 361 GAGCTTGTGATGAGATTTCAATGAGAACTTGTGTTGATGATTTGAGACCTTTC 420
DB 807 GAGCTTGTGATGAGATTTCAATGAGAACTTGTGTTGATGATTTGAGACCTTTC 866
QY 421 AACTCATATTCCCGCGCCCACTTTTCAAAATCATTGTAATGCTGAGAGTTCTTA 480
DB 867 ACTGCTCTTTTCCGCGCCCACTTTTCAAAATCATTGTAATGCTGAGAGTTCTTC 926
QY 481 AATGTCACCTTTGGGCAAAATGTTGATGATCAAGAGAGATGACCTTTGCTGAA 540
DB 927 AATGTCACCTTTGCTGCTGAGCTTTCATGACAGAGAAAGCTTGACCTTCTGCTTGA 986

QY 541 TTCTCAGAGTCCATTTGTCAAGAGGCAAGAACATGATGTTGATGACAGAAATTCAGAC 600
DB 987 TTCTCAGAGTCCATTTGTCTCAAGAGGCAAGAACATGATGTTGATGACAGAAATTCAGAC 1046
QY 601 TTGAAATCTCTTCAATGTTTGAAGAGAGAGAGATTTCTGATCCCTACCTCCT 660
DB 1047 GTTGTCTCTCTCAATGTTTGAAGAGAGAGAGATTTCTGATCCCTACCTCCTCCT 1106
QY 661 GAGACACCATGTCGCAATTTGAAACCGGTTCCAGAAATCGGTTTGAAGAGAGTTGG 720
DB 1107 GAGACCCCGTACTCCAGATTTGAGAGCAAAATTTCCAGAGATTCGGGCTCGAGCGGGGTTG 1166
QY 721 GGTGACACCGCAAGACGCGTCTGAGATGATTCATCTTTGATCTTCTTGAAGCA 780
DB 1167 GGTGACACCGCTGAGCGGCTCTGAGATGATTCAGATCTCTTGTGATCTCTTGAAGCT 1226
QY 781 ACTGATCTTGAACCCCTTGAAGAGTTCTTGAAGATTCGCCATGAGTTCAATGTTGTG 840
DB 1227 CCCGACCCGTCATCTCGAAGAGTTCTTGAAGAGTTCTTGAAGATTCGCCATGAGTTGTG 1286
QY 841 ATTCTCACTCCCAAGATTTCTGCTCAAGACATGTTTGGGATTCGCCACCGGT 900
DB 1287 ATCATGCTCCCAAGATTTCTGCTCAAGACATGTTTGGGATTCGCCACCGGT 1346
QY 901 GGGCAGGTTGTTTCAATCTTGAATCAAGTCCGAGCTTGAAGATGATGCTCTCGT 960
DB 1347 GGCAGGTTGTTTCAATCTTGAATCAAGTCCGAGCTTGAAGATGATGCTCTCGT 1406
QY 961 ATTAAGAAAGAGACATCAATCAATCCGATGATTCCTTCTTCTTCTCTCT 1020
DB 1407 ATTAAGAAAGAGACATCAATCAATCCGATGATTCCTTCTTCTCTCTCTCTCT 1466
QY 1021 GATGCTGTCGAAACAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1080
DB 1467 GATGCTGTCGAAACAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1526
QY 1081 GATGCTGTCGAAACAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1140
DB 1527 GATGCTGTCGAAACAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1586
QY 1141 TTGAAAGAAAGATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1200
DB 1587 TTGAAAGAAAGATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1643
QY 1201 AAGAGTTGACGCGCAAGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1260
DB 1644 GAGAGTTGACGCGCAAGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1703
QY 1261 GCTGCTGCTGCAATTAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1320
DB 1704 GCTGCTGCTGCAATTAATGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1763
QY 1321 AAGAGAAATATCAGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1380
DB 1764 AAGAGAAATATCAGATGATTCCTTGAAGATGATTCCTTGAAGATGATTCCT 1823
QY 1381 TCTTGCAATTTACAGTGAATTTTGAAGATGATTCCTTGAAGATGATTCCT 1440
DB 1824 TCTTGCAATTTACAGTGAATTTTGAAGATGATTCCTTGAAGATGATTCCT 1883
QY 1441 ACTTTCAGAGAAATGAG 1500
DB 1884 ACTTTCAGAGAAATGAG 1943
QY 1501 TTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1944 TTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2003
QY 1561 AACTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 2004 AACTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2063

```

QY 1621 AGTTGAAGCATTTCCATCTGAGATCGAGACCTTCTTACACCAAGTTGAGATGAA 1680
DB 2064 CGGTTGAATCTCTTCCACCTGAGATCGAGAACTCTTCCAGGATGTTGAGAACAG 2123
QY 1681 GAACACTTATGTGTGCTCATGACCGCAACAGCCAACTTCTGTTCAATGCCAAGCTT 1740
DB 2124 GAACACTTGTGTGTGTGTAAGATGAAAGAGCTTATTTTCAACATGGCAAGGCTG 2183
QY 1741 GATGTGTGCAAGAACTTAAACCGGACTGCTGAGTGTGCGGCAAGAACCAAGTTGCGT 1800
DB 2184 GACGGTGTCAAGAACTTACAGGGCTTGTGTAGTGTATGCGAAGAACTTCAAGTTGAG 2243
QY 1801 GAGTTGCTAACTCTGATGTTGATGAGTGTGATGAGCCAAAGAACTTAAAGATTGGAA 1860
DB 2244 GAATCTGCGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2303
QY 1861 GAGAAAGCTGAATGAAAGAAATGTTTGTAGCTGATGCAAGATCACTTGAACGGCCAA 1920
DB 2304 GAGCAGTCTGAGATGAAAGAAATGTAAGACCTCATGAAAGTCAAGCTGAATGGCCAG 2363
QY 1921 TTCGATGAGATATCTCTCAATGAAACGAATCCGAATGTTGAATCTTACCGATACATT 1980
DB 2364 TTCAAGTGTGATTTCTCCAGATGAAACGGGTGAGAAATGAGAGCTTACCGCTACATC 2423
QY 1981 TGCGACAGAAAGGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
DB 2424 TGTGACAGAAAGGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2483
QY 2041 GTGAGGCAATGACTTGGCGTTTGGCAACATTCGCAACCTGTAAACGATGACCAAGCCAG 2100
DB 2484 GTTGAGGCAATGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2543
QY 2101 ATTATGTGCAATGAGAAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
DB 2544 ATCATTTGTGCAATGAGAAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2603
QY 2161 GACATCTGCTGATGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
DB 2604 GAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2663
QY 2221 TCCCAAGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
DB 2664 TCAAAAGGTGCTGATGAGAAATGTAAGAAATGTAAGAAATGTAAGAAATGTAAGAAATG 2723
QY 2281 CTATTTGACCTGTGACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
DB 2724 CTGTGTGAACCTGATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2783
QY 2341 GAGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
DB 2784 GAGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2843
QY 2401 GTTCCATTGTGCGAGAGATGAA 2422
DB 2844 GTTCTCGGGCTGTGAGATGAA 2865

```

RESULT 6
AB213815
ID AB213815 standard; DNA; 2427 BP.

XX AC AB213815;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1620.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; de.
XX OS Arabidopsis thaliana.
XX XX
XX PN MO20021655-A2.

```

PD 28-FEB-2002.
PF 24-AUG-2001; 2001WO-US026685.
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
XX (SCRI ) SCRIPES RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 1620; 577bp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array of probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
XX Sequence 2427 BP; 643 A; 517 C; 587 G; 680 T; 0 U; 0 Other;
XX
XX Query Match 56.6%; Score 1486.6; DB 6; Length 2427;
XX Best Local Similarity 76.3%; Pred. No. 0;
XX Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
XX
QY 4 GCTGAGGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 63
DB 10 GCTGAGGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69
QY 64 GCTGAGGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
DB 70 TCTGAGGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
QY 124 CTGCAACACCATTAATTTATTTAGATTTGAACTTATCTCTGAAAGAAAGAAAG 183
DB 130 TTACAGCAAAACAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
QY 184 CTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
DB 190 CTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 244 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
DB 250 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
QY 304 GTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
DB 310 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 364 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
DB 370 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
QY 424 TCATCATTTCCCGCCCAACCTTTCAAAATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
DB 430 GCGTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
QY 484 GGTCACTTTTGGGCAAAATGTTTCAATGCAAGAGAGATGCAACCTTTGCTGCAATTC 543

```

Db 490 CGTCAATTATGGGCTAAGCTCTTCCATGACAAAGAGAGTTTGCTTCCATGCTTAAGTTC 549
Qy 544 CTGAGAGTCCATTGTCCACAGGGCAAGAACATGATGTTGAATGACAGAAATTCAAGACTTG 603
Db 550 CTTGCTCTTCCACAGCCACCAAGGCAAGAACTGATGTTGAGCGAGAAATTCAGAACTTC 609
Qy 604 AATGCTCTTCAACAGTTTGTAGGAAAGCAGAGAGTATCTTTGTAACCTTACCTCCAG 663
Db 610 AACACTCTGCAACACACCTTGAGAAAGCAAGAGTATCTAGCAGAGCTTAACTCCAAA 669
Qy 664 ACACCATGTGCCGAATTCGAACACCGGTTCCAGAAATCGGTTTGGAAAGAGTGGGGT 723
Db 670 ACACGTATGAAAGATTTGAGGCCCAAGTTTGAGAGATTGGCTTTGAAAGGGAGTGGGA 729
Qy 724 GACACCGAGAAACGGTGTCTGAGATGATCCAACTCTTTTGGATCTTTTGAAGCAACT 783
Db 730 GACAAATGACAGACGGTCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 789
Qy 784 GATCTCTGACCCCTTGAGAAATGTTCTGGAGAAATCCCAATGATGATGATGATGATGAT 843
Db 790 GATCTCTGACCTCTTGAGAAATGTTCTGGAGAAATCCCAATGATGATGATGATGATGAT 849
Qy 844 CTGATCTCCCAACGAGATCTTGCTCAAGCAATGTTTGGGGTATCCCGACACCGGTGG 903
Db 850 CTCTCTCCACATGATTTACTTGTCTCAGAGCAATGTTCTTGGTTACCTGACATGATGTA 909
Qy 904 CAGGTTGTTTACATCTTGATCAAGTCCGAGTTTGAAGATGAGATGATGATGATGATGAT 963
Db 910 CAGGTTGTTTACATCTTGATCAAGTCCGAGTTTGAAGATGAGATGATGATGATGATGAT 969
Qy 964 AAGCAACAGAGATCAATCAACCCCTGGAATCCATATTTATCTAGCTTTCTCTGAT 1023
Db 970 AAGCAACAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1029
Qy 1024 GCTGTCGAGCAACATGCGGTCAACGATCTGAGAAAGTATACGGAACAGAGATCGGAT 1083
Db 1030 GCGGTGAGAACTACATGCGGTGAGACGTTCCAGAGAGTTTATGATCTGATGATGATGAT 1089
Qy 1084 ATTTCTTGAGTACCTTTCAGAACAGAAAGGAAATGTTTCGAAATGAGATCTCAAGATT 1143
Db 1090 ATTTCTTGAGTACCTTTCAGAACAGAAAGGAAATGTTTCGAAATGAGATCTCAAGATT 1149
Qy 1144 GAAAAAGTGTGGCCCTACTTGGAAACCTACACAGAGATGTTGCTCATGAAATCTCCAAA 1203
Db 1150 G---AAGTCTGCGCAATCTGAGACTTACACCGAGATGCTGCGGTGAGACTTACAAA 1206
Qy 1204 GAGTTGACGAGACCGCAGATCTGATTCAGAAACGAAACGAGAGCGGCAATATGCTGCC 1263
Db 1207 GATTTGAATGGCAAGCTGACCTTATCATTTGTTATCTACAGTGAAGAAATCTTGTGCT 1266
Qy 1264 TCTTGTCTCGACATTAATTAAGTGTACACAGTGCACATCGCCATGCTTTTGAAGAG 1323
Db 1267 TCTTATATGGCTCAACAACTTGTGTCTCAGTGTACATTTGCTCANTGCTCTTGAAGAA 1326
Qy 1324 ACAAATATCCAGATTCAGATTCATTTGGAAGAAAGCTTGAAGCAATACATTTGCTCT 1383
Db 1327 ACAAAGTACCGGATTCGATATCTACCTGAAAGAGCTTGAAGCAATACATTTGCTCA 1386
Qy 1384 TGGCAATTTACAGCTGATCTTTTTCAGATGAAACATACAGATTTTCATGATCAGAGTACT 1443
Db 1387 TGGCAATTTTCATGCGGATATTTTTCAGATGAAACATGATTTTCATGATCAGAGTACT 1446
Qy 1444 TTCCAGAAATTTGACAGAGAGAGACATGTTGTCAATAAGAGGCAACCTGCTTTC 1503
Db 1447 TTCCAGAAATTTGCTGAGAGAGAGAACTGTTGGGAGTATGAAGGCAACAGCGCTTT 1506
Qy 1504 ACTCTTCTGCTCTTACCGTGTGTATCATGTATCATGTGTGTGTATCCAAATTTCAAC 1563
Db 1507 ACTCTTCTGCTGCTTATCATGTGTGTATCATGTGTGTGTGTATCCAAATTTCAAC 1566
Qy 1564 ATTTGTTTCCCTGCTGCTATGAGATATATCTTCCCTTACACGGAAGAGAGCGGAGG 1623
Db 1567 ATTTGTTTCCCTGCTGCTATGAGATATATCTTCCCTTACACGGAAGAGAGCGTGA 1626

Qy 1624 TTGAGCATTTTCCATCTCTGATGATGAGAACCTTCTTTTACCCAAAGTTTGAGATGAAGA 1683
Db 1627 TTGACTAAGTTTCCATCTCTGATGATGAGAACCTTCTTTTACCCAAAGTTTGAGATGAAGA 1686
Qy 1684 CACTTATGTGTGCTCAATGACCCGCAACAGCAATCTGTCTTCAATGCGCAAGGCTGAT 1743
Db 1687 CACTTATGTGTGCTCAATGACCCGCAACAGCAATCTGTCTTCAATGCGCAAGGCTGAT 1746
Qy 1744 CGTGTCAAGAACTTAAACCGACTCTGTGAGTGTGCGGCAAGAACCCAAAGTTGCGTAG 1803
Db 1747 CGTGTCAAGAACTTGTGAGGCTCTTGTGATGTGTGAGAGAGAAACACCCGTTGCGTAG 1806
Qy 1804 TTGCTTAACTCTGATGTTGATGAGTGTGATGAGGCAAGAAATCTAAAGTTTGAAGAG 1863
Db 1807 CTAGCTAATCTTGTGTGTGTGAGAGGAGACAGGAGGAAAGTCAAGGCAATGAAGAG 1866
Qy 1864 AAGCTGAATAATGAAGAAATGTTTGAAGCTGATGAGCAAGTGAACCTTGAACGCGCAATTC 1923
Db 1867 AAGCTGAATAATGAAGAAATGTTTGAAGCTGATGAGCAAGTGAACCTTGAACGCGCAATTC 1926
Qy 1924 AGATGATATCATCTCAATGAAACAGAAATCGAAATGTTGAACCTTGAACGATATCTTGC 1983
Db 1927 AGGTGATCTCTCTGATGAGACCGGTTAAGAACGCTGATGATGATGATGATGATGAT 1986
Qy 1984 GACACGAAGGTGCTTTGTGATGAGCTGCAATGATGAGACCTTTGATGATGATGATGAT 2043
Db 1987 GACACGAAGGTGCTTTGTGATGAGCTGCAATGATGAGACCTTTGATGATGATGATGAT 2046
Qy 2044 GAGGCAATGATCTTGGGTTTCCAACTTGGCAACCTGTAACGCTGAGCAAGCGGAGAT 2103
Db 2047 GAGGCAATGATCTTGGGTTTCCAACTTGGCAACCTGTAACGCTGAGCAAGCGGAGAT 2106
Qy 2104 ATTTGCTGAGGAAATCTGTTTCAACATTTGATCTTACATGATGATGATGATGATGATGAT 2163
Db 2107 ATTTGCTGAGGAAATCTGTTTCAACATTTGATCTTACATGATGATGATGATGATGATGAT 2166
Qy 2164 ATACTGCTGATTTCTTTGAAAGTGTGAAGAAAGATTCATCTCACTGGATTAAGATCTCC 2223
Db 2167 ACTTGTCTGATTTCTTCAACAGTGTGAAGAGATTCATCTCACTGGATTAAGATCTCA 2226
Qy 2224 CAAGGAGCTTGAACGAATGAGAGAGATTAATCAATGAAAGTTTACGAGAGAGCTA 2283
Db 2227 AAAGAGGCTTCAAGAGATTTGAGAGAAATPACCTTGGCAATTAATTAACAGAGCTC 2286
Qy 2284 TTGACCTTGAACAGAGTGTATGATTTCTGAGAGATTTTCAACCTTGAACGCTGAG 2343
Db 2287 TTGACATTTGATGATGATGATGATTTCTGAGAGATGATGATGATGATGATGATGAT 2346
Qy 2344 AGTGTGTTTACCTTGAAGATGTTTATGCTCTTAAGTACCGTAACTGCTGAATCAATT 2403
Db 2347 GCTGCGGTTTACCTTGAAGATGTTTATGATGATGATGATGATGATGATGATGATGAT 2406
Qy 2404 CCAATGCGAGAGGA 2417
Db 2407 CTTCTTGCACAGA 2420

RESULT 7
ADA68333
ID ADA68333 standard; DNA; 2427 BP.
XX ADA68333;
XX AC
XX 20-NOV-2003 (first entry)
XX DE
XX Arabidopsis thaliana gene, SEQ ID 808.
XX KM
XX Plant; Bacterial infection; fungal infection; viral infection; ds.
XX OS
XX Arabidopsis thaliana.
XX PN
PN W02003000898-A1.

XX 03-JAN-2003.
XX 22-JUN-2001; 2001WO-IB001105.
XX 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX Claim 6; SEQ ID NO 808; 899bp; English.
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX Sequence 2427 BP; 643 A; 517 C; 587 G; 680 T; 0 U; 0 Other;
XX
XX Query Match 56.6%; Score 1486.6; DB 7; Length 2427;
XX Best Local Similarity 76.3%; Pred. No. 0;
XX Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
QY 4 GCTGAGCGTCTCTCATCTGCGGTCCAGCTCTCCGTGAGCGTTGGAGAACCTTTCTT 63
DB 10 GCTGAACTTAAGTAACCGCGGTCCAGCCAACTGAGCGTTGAAAGAAACGTTGTT 69
QY 64 GCTCAGGAGAGAGATTTTGGCCCTGCTCAAGAGTCAGGGCAAGAAAGAAAT 123
DB 70 TCTGAGAAAGAGAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 129
QY 124 CTGCAACACATCAATATTATTCTAGAGTTTGAAGCTATCCCTGAAGAGACAGAAAG 183
DB 130 TTACAAACAAACAGATCATTTGTAATTCGAAGCTTTGCGTGAACAAACCGAGAA 189
QY 184 CTCGTAATGTGCAATTTTGAAGTATTTGAAGGCTAGTCAGAAAGCAATGTGTGCT 243
DB 190 CTGGAAGGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 249
QY 244 CCATGAGGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
DB 250 CCAATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
QY 304 GTTCAAGCGCTTGTGTGAGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
DB 310 CTCATGCTCTTGTGTGAGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 364 CTTGTGTAAGAGTTCAAATGAAATCTTGTGTGTAATGATTTGAGCCCTTCAAC 423
DB 370 CTCGTGTAAGAGTTCAAATGAAATCTTGTGTGTAATGATTTGAGCCCTTCAAT 429
QY 424 TCATCATTCCTCCCGCAACTCTTCAAAATCATTTGTAATGAGTGAAGTCTTAAT 483
DB 430 GCGTATATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
QY 484 CGTCACTTTTGGCAAAATTTGTCATGACAGAGAGATGCACTTTGTGCTCAATTC 543
DB 490 CGTCACTTTTGGCAAAATTTGTCATGACAGAGAGATTTGTCATTTGCTTAAGTTC 549

QY 544 CTCAGATTCATTTGTCACAAAGGCAAGACATGATGTTGATGACAGAAATTCAGAACTTG 603
DB 550 CTTGTCTTTCACAGCCACAGGGCAAGACCTGATGTTGAGCGAGAAATTCAGAACTTC 609
QY 604 AATGCTTTCACATGTTTGAAGAAAGAGAGATCTTGTGACCTTACCTCTGAG 663
DB 610 AACACTCTGCAACACCTTGAAGAAAGAGAGATCTTGAAGAGCTTAACTCCGAA 669
QY 664 AACACATGTCGCAATTCGAAACCGGTTCCAGAAATCGGTTTGGAAAGGTTGGGGT 723
DB 670 AACACTCTGCAACACCTTGAAGAAAGAGAGATCTTGAAGAGCTTAACTCCGAA 729
QY 724 GACACCGCAAGCGGTGCTGAGATGATCAATCTTCTTGTGATCTTCTTGTAGGCACT 783
DB 730 GACAAATGACAGGTGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 784 GATCCTTGCACCTTGAAGAGTCTTGAAGAAATCCATGCTGTTCAATGTTGAT 843
DB 790 GATCCTTGCACCTTGAAGAGTCTTGAAGAAATCCATGCTGTTCAATGTTGAT 849
QY 844 CTCACTCCCAAGGATCTTCTGCAAGCAATGTTTGGGATCCCGACACCGGTGG 903
DB 850 CTCTCTCAGATGTTACTTGTCTCAGAGACAAATGTTCTTGTGTTACCTGACACTG 909
QY 904 CAGGTTGTTACATCTTGATCAAGTCGAGCTTGGAGATGATGATGATGATGATGAT 963
DB 910 CAGGTTGTTACATCTTGATCAAGTCGAGCTTGGAGATGATGATGATGATGATGAT 969
QY 964 AAGCAACAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1023
DB 970 AAGCAACAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1029
QY 1024 CCTGTGCAACATGCGGTCAAGCAATGTTGAGAAATGATTCGAAACAGACATCGAT 1083
DB 1030 GCGGTGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1089
QY 1084 ATTCTTGAGATCTTCAAGCAAGAAAGGAAATGTTGCAAAATGATCTCAAGATTT 1143
DB 1090 ATTCTTGAGATCTTCAAGCAAGAAAGGAAATGTTGCAAAATGATCTCAAGATTT 1149
QY 1144 GAAAGATCTGAGCAATCTTGAAGAACTTCAACAGAGATGTTGCTCAATGATCTCA 1203
DB 1150 G---AAGTGGGCAATCTTGAAGAACTTCAACAGAGATGTTGCTCAATGATCTCA 1206
QY 1204 GAGTTGACGCGCAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1263
DB 1207 GAATGATGAGCAAGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1266
QY 1264 TCTTGTGCGCAATTAATTAAGTGTCAACAGTCAATCGCCATGCTTGAAGAG 1323
DB 1267 TCTTGTGCGCAATTAATTAAGTGTCAACAGTCAATCGCCATGCTTGAAGAG 1326
QY 1324 ACAAAATATCCAGATTCAGATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1383
DB 1327 ACAAAATATCCAGATTCAGATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1386
QY 1384 TGCATTTACAGTGAATCTTGTGCAATGACATACAGATTTCAATCATGATCA 1443
DB 1387 TGCATTTACAGTGAATCTTGTGCAATGACATACAGATTTCAATCATGATCA 1446
QY 1444 TTTCAAGAAATTTGAAGAAAGAGACATGTTGTAATTCGAGAGCACTGCTTTC 1503
DB 1447 TTTCAAGAAATTTGAAGAAAGAGACATGTTGTAATTCGAGAGCACTGCTTTC 1506
QY 1504 ACTCTTCCTGCTCTCAACGTTGTGTAATGATGATGATGATGATGATGATGATGATGAT 1563
DB 1507 ACTCTTCCTGCTCTCAACGTTGTGTAATGATGATGATGATGATGATGATGATGATGAT 1566
QY 1564 ATTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1623
DB 1567 ATTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626

QY 1624 TTGAAGCAATTCATCTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAAGA 1683
 DB 1627 TTGACTAATGTTCTCAGATCGAGATCGAGGCTCTCTACAGGATGTTGAGAAACAAG 1686
 QY 1684 CACTATGATGCTCATGATGACCGCAACAGCAATTCCTGTTCAACAATGCAAGGCTGAT 1743
 DB 1687 CACTTATGATGCTCATGATGACCGCAACAGCAATTCCTGTTCAACAATGCAAGGCTGAT 1746
 QY 1744 CGTGTCAAGAACTTAAACCGGACTCGTCAAGTGTGCGGCAAGAACCCAAAGTTGCTGAG 1803
 DB 1747 CGTGTCAAGAACTTAAACCGGACTCGTCAAGTGTGCGGCAAGAACCCAAAGTTGCTGAG 1806
 QY 1804 TTGGCTAACCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863
 DB 1807 CTAGCTAACCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
 QY 1864 AAGGCTGAATGAAGAAATGTTGAGCTGATCGAAGTGAACCTTAAACCGGCAATTC 1923
 DB 1867 AAGGCTGAATGAAGAAATGTTGAGCTGATCGAAGTGAACCTTAAACCGGCAATTC 1926
 QY 1924 AGATGATATCATCTCAATGAATGAAGAAATCCGAAATGTTGAACCTTAAACCGGCAATTC 1983
 DB 1927 AGATGATATCTCTCTCAGATGAGACCGGGAAGAAAGGATGATGATGATGATGATGATGAT 1986
 QY 1984 GACACGAAAGGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2043
 DB 1987 GACACGAAAGGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2046
 QY 2044 GAGGCAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
 DB 2047 GAGGCTATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106
 QY 2104 ATTGTCAATGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2163
 DB 2107 ATTGTCAATGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166
 QY 2164 ATACTCGTCAATCTTGTGGAAGGATGAAGAAATGATGATGATGATGATGATGATGATGAT 2223
 DB 2167 ACTCTGCTGATCTTGTGGAAGGATGAAGAAATGATGATGATGATGATGATGATGATGAT 2226
 QY 2224 CAAGGAGCTTGAAGCAATGAGAGAGATATGATGATGATGATGATGATGATGATGATGAT 2283
 DB 2227 CAAGGAGCTTGAAGCAATGAGAGAGATATGATGATGATGATGATGATGATGATGATGAT 2286
 QY 2284 TTGACCTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2343
 DB 2287 TTGACCTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2346
 QY 2344 AGTGTGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2403
 DB 2347 GCTGTGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2406
 QY 2404 CCATTGCGAGAGA 2417
 DB 2407 CCTCTTGCAGAGA 2420
 Db -
 RESULT 8
 ID ADC68423 standard; cDNA; 2714 BP.
 XX ADC68423;
 AC
 DT 18-DEC-2003 (first entry)
 XX
 DE Lolium perenne fructan biosynthesis protein cDNA SEQ ID NO:133.
 XX
 KW Lolium perenne; Festuca arundinacea; lignin; fructan; tannin;
 XX biosynthetic pathway; plant; gene; ss.
 OS Lolium perenne.
 XX
 PN W02003040306-A2.

XX 15-MAY-2003.
 PD 07-NOV-2002; 2002MO-NZ000239.
 XX 07-NOV-2001; 2001US-0337703P.
 PF 07-NOV-2001; 2001US-0337703P.
 PR (GENE-) GENESIS RES & DEV CORP LTD.
 PA (WRIG-) WRIGHTSON SEEDS LTD.
 XX
 PI Demmer J, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M;
 XX Salsbury KM, Hall C;
 DR WPI: 2003-441544/41.
 DR P-PSDB; ADC68459.
 XX
 PT New polynucleotide encoding polypeptides from Lolium perenne or Festuca
 PT arundinacea, useful for modulating the biosynthesis of lignin, fructan or
 PT tannin in a plant.
 XX
 XX Claim 1; SEQ ID NO 133; 240bp; English.
 XX
 CC The present invention describes isolated polynucleotides (i) encoding
 CC proteins (ii) from Lolium perenne and Festuca arundinacea which are
 CC active in lignin, fructan and tannin biosynthetic pathways. Also
 CC described: (1) an isolated oligonucleotide probe or primer comprising at
 CC least 10 contiguous residues complementary to 10 contiguous residues of
 CC (1); (2) a kit comprising the oligonucleotide probe or primer; (3) a
 CC genetic construct comprising (i); (4) a transgenic plant cell comprising
 CC the genetic construct of (3); (5) a plant or its seed, fruit or progeny
 CC comprising the transgenic plant cell of (4); (6) modulating one or more
 CC of the lignin, fructan or tannin compositions of a plant; (7) producing a
 CC plant having one or more of the lignin, fructan or tannin compositions;
 CC and (8) modifying the activity of (ii) involved in a lignin, fructan or
 CC tannin biosynthetic pathway in a plant. (i) can be used for modulating
 CC the biosynthesis of lignin, fructan or tannin in a plant. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 2714 BP; 656 A; 664 C; 671 G; 723 T; 0 U; 0 Other;
 Query Match 49.7%; Score 1305.4; DB 9; Length 2714;
 Best Local Similarity 72.1%; Pred. No. 0;
 Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;
 QY 16 CTGACCTGCTCCACAGATCCCGGAGCGCTTGCTGCCACCTTCTCCTCCATCTAAC 75
 DB 57 CTGACCTGCTCCACAGATCCCGGAGCGCTTGCTGCCACCTTCTCCTCCATCTAAC 116
 QY 76 GAGATTTGGCTTGTCTCAAGATGAGGCAAGAAAGAAATTTGCAACACCAT 135
 DB 117 GAGCTCATTTGCCCTTTTCAAGATGATTCACCAAGGCAAGAAATTTGCAAGCGCCAC 176
 QY 136 CAAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAAAGAAAGCTGCTAATGCT 195
 DB 177 CAGCTGCTCACTGATGTTGAAGCTTGTGAGGCTGACAAAGAGA-----GATAAGCA 230
 QY 196 GCATTTTGAAGATGATGAAGGCTAGTCAGAAACGATGCTGTTGCTCCATGAGTTGA 255
 DB 231 CCTTTGAAGACATTCCTCCGCTGCTCAGAGACATGATGCTGCCCCATGAGTTGA 290
 QY 256 CTGCTGTTGCTCCAGAGCTTGTGTTGGAAGTCACTTGAAGTGAATGTTCAAGCCCTT 315
 DB 291 CTGCTGTTGCTCCAGAGCTTGTGTTGGAAGTCACTTGAAGTGAATGTTCAAGCCCTT 350
 QY 316 GTTGTGAGGAACTCACTGTTGCTGAGTATCTCACTTCAAGGAAAGCTTGTGATGGA 375
 DB 351 GCTGTTGAGAGCTGACAGATTTCCGAGTACTGTCATTTCAAGAAACGCTTGTGATGAG 410
 QY 376 AGTTCAATGAAATCTTGTGTTGGAATGATGATTTGAGCCCTTCAATCATCATTTCCC 435
 DB 411 CAGCCAGCAGCAAGATTTGTGCTTGAAGCTGATTTTGAAGCTTCAATGATGCTTCCCA 470
 QY 436 CGCCCACTTTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495

```
Db 471 CGTCTTCATGTCCAAAGTCCATTGGAATGGGGTGGAGTTCCTTAACCGTCACTGCT 530
Qy 496 GCAAAATGTTTCATGACAGAGAGACATGACCCCTTTGCTCGCAATTCCTGAGATTCAT 555
Db 531 TCCAAGCTGTTTCAGACAGAGAGAGCTTACCACTAGCAATTCCTGAAACCCAT 590
Qy 556 TGTCACAGAGGCGACAGATGATGTTGAATGACAGAAATTCAGAACTTGAAATGCTTCA 615
Db 591 AACCAAGAGGGGACGACATGATGCTAAACGACGAATTCAGAGCTTCGTGGGCTCCAA 650
Qy 616 CATGTTTGAAGAAAGCAGAGAGATCTTGTAACCTTACCTCTGAGACACATGTC 675
Db 651 TCAGCCCTTAGAAAGGAGAGATCTAACGACATTCCTGAAAGACACCCGCTCT 710
Qy 676 GAATTCGAAACCGGTTCCAGGAATCGTTTGAAGAAGGTGGGGTGGACACCGCAGAA 735
Db 711 GAGTTCAACCAAGGTTCCAGAGCTTGAGTGGAGAAAGGTTGGGGGACACCGCTAAG 770
Qy 736 CGCGTGTGAGATGATCCAACTCCTTTGATCTTTTGAAGGCAACTGATCCTTGACAC 795
Db 771 GGTGACAGACACATCCATCTGTTCTTGAATCTTGAAGGCCCCGATCCAGCCAGC 830
Qy 796 CTGGAAGATTCCTTGGAGAAATCCCATGCTTCAATGTTGATTCGACTCCAC 855
Db 831 TTGAGAAAGTTCCTTGAACATCAATGATGTTCAATGTTGATCTCTGCTCACAC 890
Qy 856 GATCTCTGCTCAGACAAATGTTTGGGGTATCCCGACACCGGTGGGAGGTTTTC 915
Db 891 GGATATTTGCTCAATCCAAATGTTGGGATACCTGATACGAGGCGCAAGGTTGTATC 950
Qy 916 ATCTGATCAAGTCCGAGCTTGGAGATGATGATGCTCCTCGATAAGCAACAGAA 975
Db 951 ATCTGATCAAGTCCGAGCTTGGAGATGATGATGCTCCTCGATAAGCAACAGAA 1010
Qy 976 CTCAACATCACCCCTGATCTCTATTAATAGACTTCTTCTGATGCTGTCGAA 1035
Db 1011 CTGACATACCCCTGATCTCTATTAATAGACTTCTTCTGATGCTGTCGAA 1070
Qy 1036 ACATGCGGTCAACGATGTAAGAAATACGAAACAGACACTCGGATATTCCTGAGTA 1095
Db 1071 ACATGCGGTCAACGATGTAAGAAATACGAAACAGACACTCGGATATTCCTGAGTA 1130
Qy 1096 CCCTTCAAGACAGAAAGGGAATGTTGAAATGATCTCAAGATTTGAAAGCTG 1155
Db 1131 CCATTCAGAACTGAAAGGGAATGTTGAAATGATCTCAAGATTTGAAAGCTG 1187
Qy 1156 CCATTCAGAACTGAAAGGGAATGTTGAAATGATCTCAAGATTTGAAAGCTG 1215
Db 1188 CCATTCAGAACTGAAAGGGAATGTTGAAATGATCTCAAGATTTGAAAGCTG 1247
Qy 1216 AGCCGATCTGATGATGTAAGAAATACGAAACAGACACTCGGATATTCCTGAGTA 1275
Db 1248 AAGCCGATCTGATGATGTAAGAAATACGAAACAGACACTCGGATATTCCTGAGTA 1307
Qy 1276 CATAAATAGGTGACAGAGTCAACGATGTAAGAAATGTTGAAAGCTG 1335
Db 1308 CATAAATAGGTGACAGAGTCAACGATGTAAGAAATGTTGAAAGCTG 1367
Qy 1336 GATTCAGATATCTATGTAAGAAATGTTGAAAGCTGTAAGAAATGTTGAAAGCTG 1395
Db 1368 AACTCAGATATCTATGTAAGAAATGTTGAAAGCTGTAAGAAATGTTGAAAGCTG 1427
Qy 1396 GCTGATCTTTTGAATGACATGATGTTGATGATGTTGATGATGTTGATGATGTT 1455
Db 1428 GCTGATCTTTTGAATGACATGATGTTGATGATGTTGATGATGTTGATGATGTT 1487
Qy 1456 GCAGAGACCAAGGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1515
Db 1488 GCTGAGACCAAGGACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1547
Qy 1516 CTCTACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
Db 1548 CTCTACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
Qy 1576 GGTGCTGATATGAGATATATATATATATATATATATATATATATATATATATATATAT 1635
Db 1608 GGAGCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1667
Qy 1636 CATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
Db 1668 CACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1727
Qy 1696 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
Db 1728 TTGAAGACCAAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1787
Qy 1756 TTAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
Db 1788 ATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1847
Qy 1816 GTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
Db 1848 GTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1907
Qy 1876 AAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
Db 1908 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
Qy 1936 TCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995
Db 1968 GCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2027
Qy 1996 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
Db 2028 GATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2087
Qy 2056 TCGGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
Db 2088 TGTGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2147
Qy 2116 AATCTGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
Db 2148 GTGCTGTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2207
Qy 2176 TTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
Db 2208 TTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2267
Qy 2236 GAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
Db 2268 GAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2387
Qy 2356 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
Db 2388 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2447
Qy 2416 GA 2417
Db 2448 GA 2449
RESULT 9
AD68520 standard; cDNA; 2716 BP.
AD68520;
18-DEC-2003 (first entry)
Loilium perenne fructan biosynthesis protein cDNA SEQ ID NO:12.
```

KW *Lolium perenne*; *Festuca arundinacea*; lignin; fructan; tannin;
KM biosynthetic pathway; plant; gene; ss.
XX *Lolium perenne*.
OS
XX MO2003040306-A2.
PN
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002MO-NZ000239.
XX
PR 07-NOV-2001; 2001US-0337703P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (WRIG-) WRIGHTSON SEEDS LTD.
XX
PI Demmer J, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M,
PI Salisbury KM, Hall C;
XX WPI; 2003-441544/41.
DR P-PSDB; ADC68364.
XX
PT New polynucleotide encoding polypeptides from *Lolium perenne* or *Festuca*
PT *arundinacea*, useful for modulating the biosynthesis of lignin, fructan or
PT tannin in a plant.
PS
XX Claim 1; SEQ ID NO 12; 240bp; English.
XX
CC The present invention describes isolated polynucleotides (I) encoding
CC proteins (II) from *Lolium perenne* and *Festuca arundinacea* which are
CC active in lignin, fructan and tannin biosynthetic pathways. Also
CC described: (1) an isolated oligonucleotide probe or primer comprising at
CC least 10 contiguous residues complementary to 10 contiguous residues of
CC (1); (2) a kit comprising the oligonucleotide probe or primer; (3) a
CC genetic construct comprising (1); (4) a transgenic plant cell comprising
CC the genetic construct of (3); (5) a plant or its seed, fruit or progeny
CC comprising the transgenic plant cell of (4); (6) modulating one or more
CC of the lignin, fructan or tannin compositions of a plant; (7) producing a
CC plant having one or more of the lignin, fructan or tannin compositions;
CC and (8) modifying the activity of (II) involved in a lignin, fructan or
CC tannin biosynthetic pathway in a plant. (I) can be used for modulating
CC the biosynthesis of lignin, fructan or tannin in a plant. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 2716 BP; 656 A; 666 C; 672 G; 722 T; 0 U; 0 Other;
Query Match 49.7%; Score 1305.4; DB 9; Length 2716;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;
QY 16 CTCACCTGCGTCCACAGTCTCCGATGAGCGTTGATGAGACCCCTTCTGCTCAGAGAAC 75
DB 58 CTGACTCCGCTCCACAGTCTCCGATGAGCGCTTGTGCTGACCTTCTCTCCATCTTAC 117
QY 76 GAGATTTGGCTTGGCTCTCAAGATCGAGGCAAAAGAAAAGAAATTCGACACCAT 135
DB 118 GAGTCATTTGGCTCTTTTCAAGATATGTCACGAGCAAAAGAAATCTTCAAGGCAAC 177
QY 136 CAATATTTCTAGAGTTGAAAGCTATCCCTGAAGAAACAGAAAGAAAGCTGTAATGCT 195
DB 178 CAGCTGCTCAGCTGAGTTGAAAGCTTGTGAGCTGACAAAGGAGA-----GATATGCA 231
QY 196 GCATTTTGAAGTATGAGGCTAGTGAAGGAGGATCGTGGCTCCATGAGGTGCA 255
DB 232 CCGTTGAAGACATTTCTCGTGTCTCTCAAGAGAAATGTGCTGCCCCAAGGTTGCA 291
QY 256 CTTCGCTGTTCTGCAAGGCTGCTGTTTGGAGATACATTAAGATGATGTTCAAGCCCTT 315
DB 292 CTGGGCAATCAGGCTTGAAGGCTGCTGCTGAGCTATACAGGAGTAATGTTAGGAGTTG 351
QY 316 GTTGTGAGGAACTCATGTTGCTGATATCTTCACTTCAAGAAAGAGCTGTTGATGCA 375
DB 352 GCTGTGAGGAGCTGACAGTTTCCGAGTACCTGACATTCAGAAACAGCTTGTGATGAG 411

QY 376 AGTTCAATGAAACCTTGTGTTTGGAAATGATTTGAGCCCTTCAACTCATCTATTCGCC 435
DB 412 CAGCCGACAGCAAGTTTGTGCTTGAAGCTTGATTTTGAAGCTTTCATATGCTCTCTCCCA 471
QY 436 CGCCCACTCTTCAAAATTCATTTGTAATGATGTGAGTATCCAAATGCTACCTTTCG 495
DB 472 GGTCTTCATGTCATGATGTCATTTGAAATGGGGTGAAGTCTTAAACCGTCACTGTCT 531
QY 496 GCAAAATTTGTTCCATGACAGAGAGCATGACCCCTTGTGCTGAAATTCCTCAGATCCAT 555
DB 532 TCCAGCTGTTCCAGGACAGAGAGAGCTCTACCCATCTAGTAATCTTGAAGGCCAT 591
QY 556 TGTCAAGAGGCAAGAACATGATGTTGAATGACGAATTCAGAACTTGAATGCTCTTCAA 615
DB 592 AACCAACAGGACAGCAATGATGCTTAAACGACGAATTCAGAGCTTTCGAGGCTCAAA 651
QY 616 CATGTTTGAAGAAAGAGAGATATCTTGTAACCTTACTCTCTGAGACACCATGTGCC 675
DB 652 TCAGCCCTTGAAGAAAGGAGAGATATCTAACGACATTCCTGAAAGCACCCGCTCTCT 711
QY 676 GAATTCGACACCGGTTCCAGAAATCGTTTGAAGAGGTTGGGTGACACCGCAGAA 735
DB 712 GAGTTCAACACAGGTTTCCAAAGCTTGGCTTGAAGAGGTTGGGCGACACCGCTAAG 771
QY 736 CGCGTCTCGAGATGATCCAACTCTTTTGGATCTTCTTGAAGGCACTGATCTTGAAC 795
DB 772 GGTGTACAGGACACCATCCACTTGTCTTATCTTATCTTGAAGGCCCCGTGATCCAGCCAG 831
QY 796 CTTGAGAAATTTCTTGGGAAATATCCCATGATGTTCAATGTTGATTCATCTCCAC 855
DB 832 TTGAGAAAGTTCCTTGAACATATACATATGATGTTCAATGTTTATCTGTCTCCAC 891
QY 856 GATATCTGCTGAAGCAATGTTTGGGGATATCCGACACCGGTGGGAGTTGTTTAC 915
DB 892 GGATATTTGCTCATCATCATATGTTGGATACCTGTATCTGTGGCCAGGTTGTATC 951
QY 916 ATCTTGATCAAGTCCAGCTTGTGAGATGATGATGTCCTCCGTATTAAGCAACAGAA 975
DB 952 ATCTTGATCAAGTCCAGCTTGTGAGATGATGATGTCCTCCGTATTAAGCAACAGG 1011
QY 976 CTCAACATCAACCTCGAATCTCTATTTATCTAGATCTTCTCTGATGCTGTGGAACA 1035
DB 1012 CTTGACATTAACCCCTTAATGATCTCATATGTGTCACCGGCTGTGCTGTATGCTGTGAACT 1071
QY 1036 ACATGCGGTCAAGCACTTGAAGATATGAGAAACAGAGCACTCGAATTTCTGAGTA 1095
DB 1072 ACATGCGGTCAAGCACTTGAAGATATGAGAAACAGAGCACTCGAATTTCTGAGTA 1131
QY 1096 CCCTTCAAGAACAGAAAGGAAATGTTTCGAAATGATCTCAAGATTTGAAAAAGCTTGG 1155
DB 1132 CCATTCAGAACTGAGAGGAGATCTCCGTAAGTGAATCTCTGTTTG---ATGTCGG 1188
QY 1156 CCATATCTTGAAGAACTTGAAGAGATGTTGCTATGAATTCCTCAAAAGTTGACAGGC 1215
DB 1189 CCATATCTTGAAGAACTTGAAGAGATGTTGCTATGAATTCCTCAAAAGTTGACAGGC 1248
QY 1216 ACGCAGATCTGATCATCGGAAACNACAGGAGCGGCAATATGCGCTGCTGCTGCA 1275
DB 1249 AAGCTGATCTCATTTATGGAATGATGATGATGATGATGATGATGATGATGATGATG 1308
QY 1276 CATTAATTAAGGTGACACAGTGCACATGCGCCATCTTGGAGAAAGCAAAATATCA 1335
DB 1309 CATTAATTAAGGTGACACAGTGCACATGCGCCATCTTGGAGAAAGCAAAATATCA 1368
QY 1336 GATTCAGATATCTATGGAAGAAAGCTTGAAGACAAATATCAATTTCTTGGCAATTTACA 1395
DB 1369 AACTCAGACATATATTTGGAAGAAATTTGACAGGCAAAATATCAATTTCTATGCAAGTTCACT 1428
QY 1396 GCTGATCTTTTGAAGAAAGCAATGATTTTCAATCAACGATATCTTTCAGAGAAAT 1455
DB 1429 GCTGATCTTTTGAAGAAAGCAATGATTTTCAATCAACGATATCTTTCAGAGAAAT 1488

QY 1456 GCAGAGCAAGACACCTGTTGTCATATACAGAGCCACACTGCTTTCATCTTCCGTG 1515
 DB 1489 GCTGGACCAAGAGACAGGCTGGCCAGATATAGTCTCAATTGCTTCACTCCCTCAT 1548
 QY 1516 CTCTACCGTGTGATACATGATATGATGTTGATCCCAATTCACATGTTTCCCT 1575
 DB 1549 CTCTACCGTGTGATACATGATATGATGTTGATCCCAATTCACATGTTTCCCT 1608
 QY 1576 GGTGTGATATGAGATATATCTCCCTTACACCCGAGAGAGAGGAGGATGAGATTC 1635
 DB 1609 GAGAGATATGATCTGTCTACTTCCCATACATGAACTGACAAAGGCTTACCTGCTTC 1668
 QY 1636 CATCCGATGATGAGACCTTCTTATACCAAAAGTTGAAATGAAAGACATTAATGTG 1695
 DB 1669 CACCTGAAATTTGAAGAGCTCTTATACATGATATGTTGAACTGAAACAAAGTTGTA 1728
 QY 1729 TTGAAGACAAAGAACAGCTTATCATCTTCTCAATGGCTCGTCTTGAACGATGAAAG 1788
 DB 1756 CTCAATGACCCGACCAACCAATTCCTGTTCAATGACCAAGGCTTGAATGTTCAAGAC 1755
 QY 1789 ATGACAGGCTGTGATGAGATGTTGTGTAAGATGCTCATCTGAAGATTTGGCAACCTT 1848
 DB 1816 GTAGTTGATGATGATAGGCGAAAGAAATCTAAAGATTTGAAAGAGAGGCTGAATG 1875
 QY 1849 GTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1908
 DB 1876 AAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1935
 QY 1909 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1968
 DB 1936 TCTCAATGAAACAAATGCGAAATGTTGAATCTTACCAATATCTTGGACACGAAAGT 1995
 QY 1969 GCTCAGATGAAACCGGTTGCGAATGCAAGATGTAACCGCTCATTTTGGACACCAAGG 2028
 DB 1996 GCCTTTGACAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2055
 QY 2029 GCATTTGACAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2088
 DB 2056 TCGCGTTTGCACAACTTGGCACTGTAAACGTTGACCAAGCGGATTAATGTTCCAGG 2115
 QY 2089 TGTGTTTGGCCCAAAATGCGACATGATGATGATGATGATGATGATGATGATGATGATG 2148
 DB 2116 AAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2175
 QY 2149 GTGCTGTTTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2208
 DB 2176 TTCTTTGAAAGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2235
 QY 2209 TTCTTTGAAAGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2268
 DB 2236 AAACGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2295
 QY 2269 AAAGGATTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2328
 DB 2296 GAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2355
 QY 2329 GGTGTATATGTTTCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2388
 DB 2356 CTGAGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2415
 QY 2389 CTGAGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2448
 DB 2416 GA 2417
 DB 2449 GA 2450

RESULT 10
 ID ABR98516 standard; cDNA; 2829 BP.
 XX

AC ABR98516;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE cDNA sequence encoding perennial ryegrass sucrose synthase.
 XX
 DB Perennial ryegrass; sucrose phosphatase synthase; SPS; invertase; INV;
 XX sucrose synthase; SS; sucrose:sucrose 1-fructosyltransferase; SST;
 KW sucrose transporter; ST; fructan:fructan 1-fructosyltransferase; PFT;
 KW sucrose:fructan 6-fructosyltransferase; SFT; soluble carbohydrate transport;
 KW SFT; soluble carbohydrate metabolism; plant; gene; ss.
 XX
 OS Lolium perenne.
 XX
 PN WO20021130-A1.
 XX
 PD 18-APR-2002.
 XX
 PF 10-OCT-2001; 2001WO-AU001275.
 XX
 PR 10-OCT-2000; 2000AU-00000673.
 XX
 PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
 XX (AGRE-) AGRISearch LTD.
 PI Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
 DR WPI: 2002-44177/47.
 DR P-PSDB; ABR69063.
 XX
 PT New enzymes and transporter proteins involved in metabolism and/or
 PT transport of soluble carbohydrates, useful as molecular genetic markers,
 PT and in modifying soluble carbohydrate metabolism and/or transport in
 PT plants.
 XX
 PS Claim 5; Fig 74; 177bp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding perennial ryegrass enzymes or transporter proteins
 CC selected from sucrose phosphatase synthase (SPS), invertase (INV), sucrose
 CC synthase (SS), sucrose transporter (ST), sucrose:sucrose 1-
 CC fructosyltransferase (SST), fructan:fructan 1-fructosyltransferase (PFT),
 CC and sucrose:fructan 6-fructosyltransferase (SFT). The nucleic acid
 CC sequences of the invention are useful as molecular genetic markers, to
 CC isolate cDNAs and genes encoding homologous proteins from the same or
 CC other plant species, and for the modification of soluble carbohydrate
 CC metabolism and/or transport in plants. The polypeptide sequences may be
 CC used to immunise animals to produce antibodies with specificity for the
 CC proteins, and these antibodies may then be used to screen cDNA expression
 CC libraries to isolate full-length cDNA clones of interest. ABR98425-
 CC ABR98516 represent nucleotide sequences encoding perennial ryegrass
 CC enzymes or transporter proteins involved in the metabolism and/or
 CC transport of soluble carbohydrates
 XX
 SQ Sequence 2829 BP; 697 A; 678 C; 697 G; 757 T; 0 U; 0 Other;

Query Match 49.7%; Score 1305.4; DB 6; Length 2829;
 Best Local Similarity 72.1%; Pred. No. 0;
 Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;

QY 16 CTCACGCGCGCCACAGTCTCCGTGAGGTTGATGAGACCTTCTTGTGCAAGAGAC 75
 DB 111 CTGACTGCTTCCACAGTCTCCGAGGCGCTTGTGCACTTCTCTCCACCTTAC 170
 QY 76 GAGATTTTGGCTTCTCAAGATGAGGCAAGGAAAGAAATTTGCAACACCAT 135
 DB 171 GAGCTCATGTCCTCTTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 230
 QY 136 CAATATATCTAGAGTTGAAAGCTATCCCTGAAGAGAAAGAAAGAGCTGATATGAT 195
 DB 231 CAGCTGCTCACTGATGATGAGGCTTGTGATGAGCTGCAAAAGAGAGAGAGAGAGAGAG 284
 QY 196 GCATTTTGAAGATGAGAGCTAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255

Db	285	CCCTTTAAAGACATTAATCCGTGCTCTCAGAGAGCAATTGTCTGCCCCCATAGGTTTCA	344
Oy	256	CTTGCTGTGCTCCAAAGCGCTGGTGTGTGGAGTACATTAGATGATATGTTCAAGCCCTT	315
Db	345	CTGGCCATCAAGGCTTAGGCCCAGGTGTCTGGGACTTAATACGGGTGAAATGTAGCAGTTG	404
Oy	316	GTTTGTAGAGAACTCACTGTGTGAGTATCTCCACTTCAAGAAAGCTTGTATATGA	375
Db	405	GCTGTGAGGAGCTGACAGTTTCCGAGTACTTAGCAATTAAGAAACAGCTTGTATATAG	464
Oy	376	AGTTCAATGAGAACTTGTTTGSAATGSAATTTTGAAGCCCTTCAACTCATTTCC	435
Db	465	CAGCCAGCAAGAGTTGTGTGCTTGAAGTTGATTTTGAAGCTTTTCAATGCTTCTTCCG	524
Oy	436	CGCCCAACTCTTTTCAAAATCCATTGGTAATGTGTGAGTTCTTAATAGTCACTTTTG	495
Db	525	CGTCTTCAATGTCAAAGTCAATTGGAAATGGGGTCAAGTCTTAAACGTCACTGTCT	584
Oy	496	GCAAAATTTGTTCCATGACAAAGAGAGATGACCCCTTGTCTGAATTTCTCAGAGTCAT	555
Db	585	TCCAAAGCTGTCCAGGCAAGAGAGAGCTTACCCCTACTAGCACTTCTGAAACCCAT	644
Oy	556	TGTCAAGAGGCAAGAACATGATGTGAATGACAGAAATTCAGAACTTGAAATGCTTTGAA	615
Db	645	AACTAAGAGGCAACGACAAATGATGCTAAATGACAAATTCAGACCTTGTGAGGCTCAG	704
Oy	616	CATGTTTTGAGAAAGCAGAGAGATATCTTGATCCCTACTCTTGACACATAGTCC	675
Db	705	TCAGCCCTTGAAGAAAGCAGAAAGATTTTAACAGAGATCCCTGAGAGACACCCGCTCT	764
Oy	676	GAATTCGAACACCGGTTCCAGAAATCGGTTGAAAAGTTGGGGTGACACCGCAGAA	735
Db	765	GAGTTCAACCAAGGTTCCAAAGAGCTGTGCTTGAGAAAGGTTGGGGCCACACCCCTAAG	824
Oy	736	CGCGTGTCAAGATGATCCAACTCCCTTTGSAATCTTGAAGCAATGATCCTTTGAC	795
Db	825	CGGTACAGAGCTCCATCCACTTGCTTCTTGATCTTGAAGGCCCTGTATCAGCCAGC	884
Oy	796	CTTGAGAAGTTCCCTGGGAGAAATCCCATAGTGTTCATATGTTGTGATTTCTCACTCCAC	855
Db	885	TTGGAGAAAGTTCTTGGAACTTAACAAATGATTTTCAATGTTTATCTGTCTCCACAC	944
Oy	856	GGAATCTTGCCTCAAGCAATGTTTTGGGGTATCCCGACACCGGTGGCAGGTTGTTAC	915
Db	945	GGATATTTTCTCAATCAATGATGTGTGGGATACCTGATATCGTGGCCAGGTTGTATAC	1004
Oy	916	ATCTTGGATCAATGCCAGCTTTGGAGATGAGATGCTCTCCGTATTAAGAACAAAGA	975
Db	1005	ATCTTGGATCAAGTCCGTGCTTGTGGAGATGAGATCTTCTAGGAATTAAGCAGCAAGC	1064
Oy	976	CTCAACATCAACCCCTGAAATCTCATTTATTAATAGACTTCTCGATAGCTGTGGAGCA	1035
Db	1065	CTTGACATTAACCCCTTAAGATCTCTCATTTGTTCACAGGCTGTGCTGTATGCTGTGAAC	1124
Oy	1036	ACATCGGTCAACGACTTGAAGAAATTAACGAAACAGACACTCGGATATTTCTTGAGTA	1095
Db	1125	ACATGTGGCCAGGGCTGAGAAAGTTAATGGGACTGAGCACTGACATTTCTCGGTGT	1184
Oy	1096	CCCTTCAGAAACAGAAAGGAAATTTTCCAAAATGATCTCAAGATTTGAAAAAGTCTGG	1155
Db	1185	CCATTCAGAACTGAGAGGGGATCTCCGTAAGTGAATCTTCGTTTTG--ATGTCGG	1241
Oy	1156	CCATTACTTGGAAACTTAACAAGAGATGTTGCTCATGAAATCTCCAAAAGAGTTGACGGC	1215
Db	1242	CCATTACTTGGAGATTAACCGAGAGATGTTGCAAAAGAACTATAGAGAAAAATGACACC	1301
Oy	1216	ACGCGCAAGATCTGATTCATCGGAAACNAACGACCGGCAATATGCTGCTCTTGTCTGCA	1275
Db	1302	AAGCTGATCTCATTAATGGAACACTACAGACGGTAACCTTGTGGCCACTGTGTTGG	1361
Oy	1276	CATTAATTAAGGTGACACAGTGACCAATGCGCCATGCTTTTGAAGAACAAATATATCA	1335
Db	1362	CATTAAGTTGAGGTTTAAACCAAGTGACCACTTGGCCATGCTTTGAGAAAACAAAGTATCCA	1421
Oy	1336	GATTCAATATCTAATTGAGAAAGCTTTGAAGAACAAATATCAATTTCTTGTGCAATTTACA	1395
Db	1422	AACCTCAATATTAATTGGAACAAATTTGACAGCCAAATCAATTTCTCATGCGCAATTA	1481
Oy	1396	GCTGATCTTTTGAATGAATCAATCAATTTCAATCAATCAACCAATCTTTCAGGAATTT	1455
Db	1482	GCTGACCTTAATTTGCAATGAATATACATGAATTTCAATCAATCAACCAATCTTTCAGGAATTT	1541
Oy	1456	GCAGAGAACAGAGACACTTGTGTAATACAGAGACCAACTGTCTTCACTCTCTGT	1515
Db	1542	GCTGAGAACAGAGACAGCTGGGCCAGTATGATGTCACATTTGCTTCAACCTCCGTAT	1601
Oy	1516	CTCTACCTGTTGTATCATGATATGATGTGTTGATATCCAAATTCACATTTGTTCCCT	1575
Db	1602	CTCTACCTGTTGTATCATGATATGATGTGTTGATATCCAAATTCACATTTGTTCCCT	1661
Oy	1576	GGTCTGAATAGGAATATATCTTCCCTTACACCGAAGAAACGAGAGTTGAACATTTCC	1635
Db	1662	GGAGCAGATATGATGTCATCTTCCATATCACTGAACCTGACAAAGGCTTACGTCCCTC	1721
Oy	1636	CATCTGAGATGSAAGACCTTCTTATACCAAAAGTTGAGATGAAGAACATTAATGTGTG	1695
Db	1722	CACCTGAAATTTGAAGAGCTCTTTTACAGCATGTGGAACCTGAAACAAATTTGTA	1781
Oy	1696	CTCAATGACCGCAACAGCAATTCCTGTTCACAAATGCCAAGGCTTGAATCGTGTCAAGAC	1755
Db	1782	TTGAAGAACAGAACAGCTATCATCTTTCATATGCTGTCTTGAACCGTGTGAAGAC	1841
Oy	1756	TTTACCGGACTGTGAGATGTGTGGGCAAGAACCCAAAGTTGCGTGAATGTGCTTACCTC	1815
Db	1842	ATGACAGGCTTGTGAGATGTGCTGTGAAGATGTCTATCGAAGATTTGGCAAACTT	1901
Oy	1816	GTAATGTATGATGTGATATGAGGAAAGAAATCTAAAGTTTGAAGAGAGGCTGAATG	1875
Db	1902	GTAATGTGTGCTGTGTGACATGACAGAGATTCAGAGATAGGAGAGACGCTGAATTC	1961
Oy	1876	AAGAAATGTTTGAAGCTGATGACAAAGTACAAATTTGAACGGCCAAATTCAGATGATATCA	1935
Db	1962	AAAAGATGTACAGTCAATGATGAGAGATTAAGTTGAGGGCCAAATTCGGGTGATCTCA	2021
Oy	1936	TCTCAATGAACAGAAATCCGAATTTTGAACCTTTACGATATCATTTGGCAGACGAAAGCT	1995
Db	2022	GCTCAAGATAACCGGTTTCGCAATGCAAGTTGTAACCGCTTCAATTTGTGACACCAAGGCG	2081
Oy	1996	GCCTTTGTACAGCCGCAATTTGATGAGACCTTTGGAATGAAGTTGTGAGGGCAATGACT	2055
Db	2082	GCAATTTGTACAGCTGTCAATCTTATGAGAACATTTGGCTGACTGTGTGTGAGGCAATACA	2141
Oy	2056	TGCGGTTTGCACAACTTGCAACTGTATACGAGTGAACCAAGCCGAGATTAATTTGTCAGAGG	2115
Db	2142	TGTGTGTTGCCCAATATGACAGATGCGCATGTGTGGCCAGCTGAATATCTGTGAATGCT	2201
Oy	2116	AAATCTGTTTCAACATTTGATCTTATCCATGATGATCAAGTCTGTGACATATCTGTGAT	2175
Db	2202	GCTCTGTTTGTCAATGATCTTATACCAAGTGAACAAGCTGTGAGATATCTGTGTCAAC	2261

OY 2416 GA 2417
DB 2502 GA 2503

RESULT 11
AAC66090
ID AAC66090 standard; cDNA; 2494 BP.

AC AAC66090;
AC AAC66090;
DT 13-FEB-2001 (first entry)
XX
DE Rice sucrose synthase cDNA sequence.
XX Cold resistance; transgenic plant; sucrose decomposition; rice;
KM sucrose synthase; 89.
XX
OS Oryza sativa.
XX
FN JF2000245279-A.
XX
PD 12-SEP-2000.
XX
PP 01-MAR-1999; 99JP-00052102.
XX
PR 01-MAR-1999; 99JP-00052102.
XX
PA (MITA) MITSUI CHEM INC.
XX
DR MPI: 2000-675173/66.
XX P-PSDB; AAY85664.

PT Novel method for the reinforcement of cold resistance in a plant
PT comprising introducing a vector encoding an enzyme that decomposes
PT sucrose into the plant.
XX
PS Claim 1; Page 8-11; 22pp; Japanese.

CC This invention relates to a method for the reinforcement of cold
CC resistance in a plant, comprising introducing an expression vector having
CC a DNA encoding an enzyme for decomposing sucrose connected downstream to
CC a promoter expressible in the plant, and expressing the enzyme in the
CC plant body. Included in the invention are an expression vector used in
CC the method, a transformed plant carrying the expression vector, and a
CC transformed rice carrying the above expression vector. The method is used
CC for reinforcing cold resistance in a plant. The present sequence
CC represents cDNA encoding sucrose synthase used in the method

XX
XX Sequence 2494 BP; 622 A; 599 C; 628 G; 645 T; 0 U; 0 Other;
SQ

Query Match 49.5%; Score 1300.6; DB 3; Length 2494;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 1728; Conservative 0; Mismatches 665; Indels 9; Gaps 2;

OY 16 CTCACGCGGTCACAGTCTCCGTGAGCGTTGATGAGACCTTCTTCTCAGAGAAC 75
DB CTACTTCGCTCCACAGTCTCCGCGAAGCGCTCGGTGCCACCTTCTCTCATCCCAAT 95
OY 36 CTACTTCGCTCCACAGTCTCCGCGAAGCGCTCGGTGCCACCTTCTCTCATCCCAAT 95
DB CTACTTCGCTCCACAGTCTCCGCGAAGCGCTCGGTGCCACCTTCTCTCATCCCAAT 95
OY 76 GAGATTTTGGCGCTGCTCAAGATGAGGCGAAGGAAAGAAATTCGACACACAT 135
DB GAGATTTTGGCGCTGCTCAAGATGAGGCGAAGGAAAGAAATTCGACACACAT 135
OY 96 GAGATTTTGGCGCTGCTCAAGATGAGGCGAAGGAAAGAAATTCGACACACAT 155
DB GAGATTTTGGCGCTGCTCAAGATGAGGCGAAGGAAAGAAATTCGACACACAT 155
OY 136 CAATTTATCTAGAGTTGAAGCTATCCCTGAGAGAAAGAAAGAGCTGCTAATGAT 195
DB CAGCTGCTTGGGAGGAGTTGAGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 209
OY 156 CAGCTGCTTGGGAGGAGTTGAGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 209
DB CAGCTGCTTGGGAGGAGTTGAGAGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 209
OY 196 GCATTTTGAAGTATGAGAGCTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255
DB CCGCTTGAAGAGATCTCCGAGGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 269
OY 210 CCGCTTGAAGAGATCTCCGAGGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 269
DB CCGCTTGAAGAGATCTCCGAGGCTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 269
OY 256 CTGCTGTTGCTGCAAGGCGTGGTGTGGAGATAGATTAAGTGAATGTTCAAGCCCTT 315
DB CTGCTGTTGCTGCAAGGCGTGGTGTGGAGATAGATTAAGTGAATGTTCAAGCCCTT 315

DB 270 CTGCGCATCAAGGCCGAGGCTGGTGTCTGGGAGATACATTCGGGAGTAATGTAAGTAGTGG 329
OY 316 GTTGTGAGAGAACTCATCTGTGCTGAGTATCTCCACTTCAAGAGAGAGCTTGTGAGAGA 375
DB 330 GAGGTGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 389
OY 376 AGTTCAATGAGAACTTGTGTTTGAATGAGATTTGAGCCCTTCAATCTCATATTCCTCC 435
DB CAGCAG 449
OY 436 CGCCCACTCTTCAAAATTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 495
DB CGCCCACTCTTCAAAATTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 509
OY 496 GCAAAATTTGTTCCATGACAG 555
DB TCCAGATTTGTTCCATGACAG 569
OY 556 TGTCAAGAGGAG 615
DB AACCAAGAGGAG 629
OY 616 CATGTTTGAAG 675
DB TCATCCCTTGAAG 689
OY 676 GAATTCGAGACACCGGTTCCAGAGAAATGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 735
DB GAGTTCAACACAGAGTTTCCAG 749
OY 736 CGGAGTCTGAG 795
DB CGGAGTCTGAG 809
OY 796 CTTGAGAGAGTTCTTGGAG 855
DB TTGAGAGAGTTCTTGGAG 869
OY 856 GAGATCTTCTGAG 915
DB GAGATCTTCTGAG 929
OY 916 ATCTTGATGAG 975
DB ATCTTGATGAG 989
OY 976 CTCAACATCAACCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGAT 1035
DB CTCAACATCAACCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGATCTCTGAT 1049
OY 1036 ACATGCGGTCACAGACTTGAAGAGATATACGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1095
DB ACATGCGGTCACAGACTTGAAGAGATATACGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1109
OY 1096 CCGTTCAAG 1155
DB CCGTTCAAG 1166
OY 1156 CCATATCTTGAAG 1215
DB CCATATCTTGAAG 1226
OY 1216 AGCGAGATCTGATCATCGAG 1275
DB AGCGAGATCTGATCATCGAG 1286
OY 1276 CATAAATTAAGAGTCAACAG 1335
DB CATAAATTAAGAGTCAACAG 1346
OY 1336 GATTCAATATCTATTTGAG 1395
DB GATTCAATATCTATTTGAG 1395

|||||
Db 2458 ACTGCTCCTACCTTGAGATGCTACGCTCTCAAGTACCGCAAAATGGTACACCGTT 2517
QY 2404 CCATTGGC 2411
Db 2518 CCATTGGC 2525
RESULT 13
AD68521
ID AD68521 standard; cDNA; 2754 BP.
XX
AC AD68521;
XX
DT 18-DEC-2003 (first entry)
DE S. arundinaceus fructan biosynthesis protein cDNA SEQ ID NO:13.
XX
KM Lolium perenne; Festuca arundinacea; lignin; fructan; tannin;
XX biosynthetic pathway; plant; gene; ss.
XX
OS Schedonorus arundinaceus.
XX
PN WO2003040306-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-NZ000239.
XX
PR 07-NOV-2001; 2001US-0337703P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX (WRIG-) WRIGHSON SEEDS LTD.
PI Demmer J, Forster RL, Gibson JB, Shenk MA, Norries MG, Glenn M;
PI Saulebury KM, Hall C;
XX
DR WPI: 2003-441544/41.
DR P-PSDB; AD68365.
XX
PT New polynucleotide encoding polypeptides from Lolium perenne or Festuca
PT arundinacea, useful for modulating the biosynthesis of lignin, fructan or
PT tannin in a plant.
XX
XX Claim 1; SEQ ID NO 13; 240bp; English.
XX
PS The present invention describes isolated polynucleotides (I) encoding
CC proteins (II) from Lolium perenne and Festuca arundinacea which are
CC active in lignin, fructan and tannin biosynthetic pathways. Also
CC described: (1) an isolated oligonucleotide probe or primer comprising at
CC least 10 contiguous residues complementary to 10 contiguous residues of
CC (I); (2) a kit comprising the oligonucleotide probe or primer; (3) a
CC genetic construct comprising (I); (4) a transgenic plant cell comprising
CC the genetic construct of (3); (5) a plant or its seed, fruit or progeny
CC comprising the transgenic plant cell of (4); (6) modulating one or more
CC of the lignin, fructan or tannin compositions of a plant; (7) producing a
CC plant having one or more of the lignin, fructan or tannin compositions;
CC and (8) modifying the activity of (II) involved in a lignin, fructan or
CC tannin biosynthetic pathway in a plant. (I) can be used for modulating
CC the biosynthesis of lignin, fructan or tannin in a plant. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 2754 BP; 652 A; 685 C; 684 G; 733 T; 0 U; 0 Other;
Query Match 49.3%; Score 1295.4; DB 9; Length 2754;
Best Local Similarity 70.2%; Pred. No. 0; Mismatches 742; Indels 9; Gaps 2;
Matches 1769; Conservative 0; Mismatches 742; Indels 9; Gaps 2;
QY 16 CTCACCTGCGCTCAACAGTCTCCGTGAGCGTTGGATGAGACCTCTTCTGTCAAGAGAAC 75
Db 125 CTGACTGCTCTCAACAGTCTCCGGAGCGCTTGGTGCACCTTCTCTCCATCCCAAC 184
QY 76 GAGATTTGGCCTTGCTCTCAAGATCGAGGGCAAGAAAGAAATTTCTGCAACACCAT 135

|||||
Db 185 GAGCTCAATGCGCTCTTTTCCAAAGTATGTTTCCACGAGGCAAAAGAAAGCTTCACGGCAC 244
QY 136 CAAATTTATTCAGAGTTTGAAGCTATCCCTGAAGAGAAACAGAAAGAGCTCGTAAATGCT 195
Db 245 CAGCTGCTCACTAGAGTTTGAAGCATTTGTTGAGTCTGACAAAGAGAGG-----TATGCA 298
QY 196 GCATTTTGAAGTATGAAGGCTAGTCAGAGAGCGATCGTGTTCCTCCATGGGTTGCA 255
Db 299 CTTTTCAAAGACATTCCTCCGTGCTGCCAGAAAGCAATTTGCTTCCCTCATGGGTTGCA 358
QY 256 CTTGCTGTGTGCTCCAAAGGCGTGTGTTTGGAGTACATTGAGTAAATGTTCAAGCCCTT 315
Db 359 CTTGCCATCAAGGCCAGGCTGTGTCTGGACTACATTAAGGGTGAATGTTAGAGAGTTG 418
QY 316 GTTGTGAGGAACTCACTGTGTGAGTATCTCCACTTCAAGAGAGAGCTTGTATGAGA 375
Db 419 GCGGTGAGAGAGCTGACAGATTTCCGAGTACTAGACATTCAAGAGAGAGAGCTTGTATGAG 478
QY 376 AGTTCAAAATGAAAATTGTTTGGAAATGGATTTGAGCCCTTCAACTCATATTCGCC 435
Db 479 CAGCCGACAGCAAGTTTGTGCTTGAAGCTTGAATTTGAGCCTTCAATGCTCCTCCCA 538
QY 436 CGCCCACTCTTCAAAAATCCATTGTAATGCTGTGAGTTCCTAATTCGCACTTCG 495
Db 539 CGTCTTCATGTCAAAGTCATTCATTTGAATGGGGTGAAGTTCCTTAAACCGTCACTGTCT 598
QY 496 GCAAAATTTGTTCCATGACAGAGAGAGATGACCCCTTGTGCTGCAATTCCTCAGAGTCAT 555
Db 599 TCCAGCTGTTCCAGACAGAGAGAGAGGCTCTACCCACTACTGAACTTCCTGAAAGCCCAT 658
QY 556 TGTCACAAGGCAAGAAATGATGTTGAATGACAAATTCAGAACTTGAATGCTCTTCA 615
Db 659 AACCAAGAGGCAAGCAATGATGCTAAACGACGAATTCAGAGCTTCGGGGCTCCAA 718
QY 616 CAGTTTGAAGAAAGAGAGAGTATCTGTACCTTACTCCTGAGACACATGAGCC 675
Db 719 TCAGCCCTTGAAGAAAGAGAGATATCTAACAAGATCTTGAAGACACCCGCTCTTT 778
QY 676 GAATTCGAAACCGGTTCCAGAGAAATGCTTGGAAAGAGTTGGGGTGAACCGCAGAA 735
Db 779 GAGTTCAACACAGAGTTTCAAGAGCTTGGCTTGAAGAGGGTTGGGTTGACACCGCTAAG 838
QY 736 CGCGTCTCGAATGATTCACATCTCTTTGAGATCTTTGAGGCACTGATCCTTGCAAC 795
Db 839 CGTGTACAGACACCAATCTGCTTGTGATCTGTAGAGCCCTGATCCGACAGC 898
QY 796 CTTGAGAAAGTTCCCTGGGAGAAATCCCATGATGTTCAATGTGATTCACCTCCAC 855
Db 899 TTGAGAGAGTTCTTGAACCTATACGATGATGTTCAATGTCGTTATCTGTCACAC 958
QY 856 GGAATCTTCGCTCAAGCAATGTTTGGGGTATCCGACACCGGTGGCAGGTTGTTTAC 915
Db 959 GGAATTTTGTCTCAATCAATGATGTTGGATACCTGATACCGGTGGCAGGTTGTATAC 1018
QY 916 ATCTTGAATCAAGTCCGAGCTTTGAGAGATGAGATGCTCTCTCGTATPAAAGCAACANGA 975
Db 1019 ATCTTGAATCAAGTCCGCTTGTGAGAAATGAGATGCTTGTGAGAGATTAAGACAGCAAGGC 1078
QY 976 CTCAACATCAACCCCTGAAATCCATATTAATTAAGATTTCTTCTGATGCTGTGAGAA 1035
Db 1079 CTTGACATTAACCCCTTAAGATTCATATGTCAACCAAGTTGTGCTGTAGTGTGAGACT 1138
QY 1036 ACATTCGCTCAACGACTTGAAGAAATGATPACGSAACAGAGACTCGATATTTCTGAGTA 1095
Db 1139 ACATGTGCGCAGCGCTGAGAGAGTTATTTGGAGATGAGACACTGACATTTCTTCGTTT 1198
QY 1096 CCTTTCAGAAACAGAAAGGAATTTGTGAATAATGATTCGAAGATTTGAAAAAGTCTGG 1155
Db 1199 CCAATTCAAGAACTGAAAAAGGGAATCTCCGTAAGTGAATATCTGTTTGG--ATGTCGG 1255
QY 1156 CCAATCTTGAAGAACTPACAGAGAGATGTTGCTCATTAATTTCCAAAGAGTTGCAAGCGC 1215


```
Db 19 GACGCGCTCGAGCGCCCTCCACAGCGTCAGGAGCGCATCGCGACTCCCTCTCCGCC 78
Qy 67 CACAGSACGAGATTTTGGCTTGTCTCAAGATCGAGGGCAAGAAAAGAAATTCG 126
Db 79 CACCCCAATGAGTGTGTGCGCGTCTTCAACAGGCTGTTAACCTCGGAAAGGAATGTT 138
Qy 127 CAACACCATCAATTTATTTCTAGAGT--TTAGAGTATCCTGAAAGAGAACAAAGAG 183
Db 139 CAGGCGCCACAGATCATTTGATGAGCAACAAACGAATCTCGAGGCTGATCGTAGAAG 198
Qy 184 CTGCTATAGTGATTTTGAAGTATGAAAGGTATCTAGSAAAGSAGTGTGTGCT 243
Db 199 CTGAAGACGGGTCTTTGAGAGATGCTGAGAGAGTGCACAGAGGGGATTTATCTCC 258
Qy 244 CCATGGGTGACATTTGCTGTTCTGTCAGAGCTGGTGTGTTGGAGTACATTTAGAGT 303
Db 259 CCAATGGTGTGCTTTGCCATCGGCCGAGGCCGGTGTCTGGAGTATGTAGGGTCAAT 318
Qy 304 GTTACGCGCTTGTGTTGAGGAACTCATGTGTGATCTCCACTTCAAGAAAG 363
Db 319 GTGACGAGCTCGCTGTGATGTGCTGACAGTCCCGAGTACTTGCAAGTTCAAGAAAG 378
Qy 364 CTGTGTATGTAATTTCAATGSAATTTGTTGAAATTTGAAATTTGACCCCTTCAAC 423
Db 379 CTGTGGAAGAGGACCAACAACTTTGTGCTTGAGCTGAGTTTCAGAGCCATTCAT 438
Qy 424 TCATCATTTCCCGCCCAACTCTTTCAAATTCATTTGTAATGTTGAGTTCCTAAT 483
Db 439 GCGTCTTCCCTGCTCTCTGTGCAAGTCAATGGCAATGTGTGTCAATTTCTCTAAC 498
Qy 484 CGTCACTTTGGCAAAATTTGTTCAATGACAGAGAGCATGACCCCTTGTCTGAATTC 543
Db 499 AGGATCTGTATCGAAGCTCTTCATGACAAAGGAGACATGTACCCCTCTCAACTTC 558
Qy 544 CTCGAGTCCATTTGTCAGAGGAGCAAGATGATGATTTGAAATGACAGATTCAGAACTTG 603
Db 559 CTGTGTCACAACTACAAAGGAGCATGACAGATGTTGAATGACAGATTCGAGTCTC 618
Qy 604 AATGCTCTTCAACATTTTGAAGAAACAGAGATCTTTGGTACCTTACTCTGAG 663
Db 619 AGTGTCTCCAAAGGTCTCTGAGAGGCTGAGAGCATCTTTCTGTCTTTCAGCAGAC 678
Qy 664 ACACCATGTGCGCAATTTGAAACACCGGTTTCAGAGAAATCGTGTGAAAGAGTTGGGT 723
Db 679 ACCCTTACTCAGATTTCCACCAACAGGTTCCAGSAACTTGGTCTGAGAAAGGTTGGGT 738
Qy 724 GACACCGGAAACGCGTGTCTGAGATGATCCAACTCCCTTTTGGATCTTTGAGCAACT 783
Db 739 GACTGTGCAAGCGTTCCAGAGACCAATTCACCTCTTTTGAACCTTCTTGAAGGCCCT 798
Qy 784 GATCTTGGACCCCTTGAGAGTCTCTTGGAGAAATCCCATGATGTTCAATGTTGTAT 843
Db 799 GATCGTCCACCTCGAGAAATTTCTTGGAAACATCCCAATGTGTTCATGTTGTCTAT 858
Qy 844 CTCACCTCCCAACGAACTTGTGCTCAGACAAATGTTTGGGATTCGCGACACCGGTGG 903
Db 859 ATGTCCCAACGAGTTACTTGGCGCAAGCCATGTCTTGGGTTACCTCGACACTGGAGGG 918
Qy 904 CAGGTGTTTCACTTTGAGATCAAGTCGAGCTTTGAGATGAGAGTCTCTCCGATA 963
Db 919 CAGGTGTCTACATTTTGAACCAAGTCGTGTATGAGAAATGAGATGCTTGAAGATC 978
Qy 964 AAGCAACAGAGATCAATCAACCCCTGATCTCATTTACTATAGACTTCTCTGAT 1023
Db 979 AAGCAGAGAGGTCTCAATTTACACAGCATCTTATTTGTCACAGGTTGCTCCGTAT 1038
Qy 1024 GCTGTGGAACAACATGCGGTCAACGACTTGAAGAAATGATACGGAACAGACACTCGAT 1083
Db 1039 GCACTGGGCAACACTGCGGTGACGCTTGTGAAGAGCTTTGGACACGAGCAACCCAC 1098
Qy 1084 ATTCTGAGTACCTTTCAGAAACGAAAGGAATTTGTTCAAAATGATCTCAAGATTT 1143
Db 1099 ATCTACGTGTGCTTTCAAGAACTGAAATGGAATTTGTGCAAGTGAATCTCACTTT 1158
Qy 1144 GAAAAAGTCTGGCCATATCTTGAAGAACTTACACAGAGAGTGTGCTCATGAATCTCCAA 1203
Db 1159 G---AAGTCTGGCGGTACTGTGAAGAACTTCACTGATGATGTGGGACGAAATGTCTGGA 1215
Qy 1204 GAGTTGACGGCACGCCAGATCTGATATGGAACCAACAGCCGACGCAATATGTGGCC 1263
Db 1216 GAGTCCAGGCAATCTGACTTGAATCATGGAACATCAAGTACGGAACCTTGTGTGA 1275
Qy 1264 TCTTGTCTGACATAAATTTAGGTGACACAGGACCAATCGGCCATGTGCTTGAAG 1323
Db 1276 TGCTTGTCTGACACAGATGAGGTGTTACCATTTGTACATTTGCCATGTGCTGAGAA 1335
Qy 1324 ACAAAATATCAAGATTCATATCTATTTGAGAGAGCTTGAAGACAAATACATTTCT 1383
Db 1336 ACCAATGACCCCAACTGTGACCTTACTGGAAGAGTTGAGGATCACTACACTTCTCA 1395
Qy 1384 TGGCAATTTACAGTGTATCTTTTGCATGACATACATATTCATCATCAGTATCT 1443
Db 1396 TGGCAGTTCAACACTGACTGATTTGCGATGAACACGCTGACTTCATCATCACCAATCC 1455
Qy 1444 TTCCAGGAATTTGAGAGACAAAGACACTGTTGTCAATACGAGACCACTGCTTTC 1503
Db 1456 TTCCAGAGATTTGCCGAAACAGAGACACTGTTGTGATGATGATCTCATGTGGCATTC 1515
Qy 1504 ACTTTCCTGTGTCTACACGCTGTGTATCATGTATTCAGTGTGTGATCCCAATTCAC 1563
Db 1516 ACCATGCTGTGTGTATCCGATGTTGTCACGATATGATGTTTGTGACCCGAAGTTTAC 1575
Qy 1564 ATTGTTCCTGTGTGTGTATGAGATATATCTTCCCTTACACCGAAGAGAGGAGG 1623
Db 1576 ATTGTCTGCTGTGTGTGACATGTCCATCTACTTCTTACTGTGTGTGTGTGTGTGT 1635
Qy 1624 TTGAAGATTTCCATCTGAGATGAGAAACCTTCTTACCAAAAGTTGAATGAGAA 1683
Db 1636 CTCACCTCCCTCCACCAAGATTTGAGAGTGTCTTACAGCAAGTTGACACAGAGAG 1695
Qy 1684 CACTTATGTGTCTCAATGACCGAACAGCAATCTTGTTCACATGATCCAGGCTTGT 1743
Db 1696 CACAAATTTATGTGAGAGACAGAAACAGCAATCATCTTCTCATGTGTGTGTGTGTGT 1755
Qy 1744 CGTGTCAAGAACTTAAACCGGATCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGT 1803
Db 1756 CGTGTCAAGAACTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1815
Qy 1804 TTGCTTAACTCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1863
Db 1816 CTGTGTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1875
Qy 1864 AAGGCTGAATGAGAAATGTTTGAAGCTGATGACAAATGTAACCTTGAACCGGCAATTC 1923
Db 1876 CAGGCTAGTTCAAGAAATGTTTGAAGCTTATGAGCAATTAACCTTGAATGAGGCAATTC 1935
Qy 1924 AGATGATATCATCTCAAGAAACAGAAATGGAATGTGAACTTTTCCGATACATTTGC 1983
Db 1936 CGCTGATCTCCGCGAGAGAAACCGTGTGCGAAACGTAAGCTTACCTTATCATCTGC 1995
Qy 1984 GACAGAAAGGTGCTTTGACAGCTTGAATGATGAACTTTGATGACAGTGTG 2043
Db 1996 GACACCAAGGTGCTTTTGTGACGCGCTTTCTAGAGAGATTCGGTCTCATCTGTGTT 2055
Qy 2044 GAGGCAATGACTTGCAGTTTGCACAACTTGCACACTGTAAACGCTGAGCAACGCGAAT 2103
Db 2056 GAGTCCATGACGTGTGTCTCCGACATTTGCAACCGCTATGTGTGTCCAGCTGATGATC 2115
Qy 2104 ATTGTCCATGGAATTTGTGTTTCAACATTTATCTTACATGTGTATCAAGTGTGATC 2163
Db 2116 ATGTTGAACGAGTGTCTGCTTTCACATTTGACCAATGACAGGTGACAGGCTCGGCG 2175
Qy 2164 ATACTGTGATTTCTTTGAAAGTGTAAAGATCATCTGACGATTAAGATCTCC 2223
Db 2176 CTGTCTGTGATTTCTTTGAAAGTGTGCAAGAGACCCAGCCTGTGACCAAGATCTCG 2235
```


OY 2224 CAAGAGGCTTGAACGAAATAGAGAGATATCATGAGATTCTCGAGAGACTA 2283
 DB 2236 CAGGCGCGCTCTTCAAGCTATGAGAGAAATACCTGGAAGCTTACTCTGAAGGCTG 2255
 OY 2284 TTGACCTCGACAGAGATGATGATTTCTGGAAGCATGTTTCCAACTTGAAAGCGCGTGA 2343
 DB 2296 ATGACCCCTCAACCGCGCTTACGATTTCTGGAAGTACGTTCTGGAAGCTGGAAGCGTGA 2355
 OY 2344 AGTGTGCTTACCTTGAAGATTTTATGCTCTTAAGTACCGTAAGCTGCGTGAATCACTT 2403
 DB 2356 ACCGCGCGCTCTTGAAGATCTGTATGCTTCAAGTACCGCAACATGCGTGAAGCGCTC 2415
 OY 2404 CCATTGCGAGAGAG 2418
 DB 2416 CCGTTGCGCTGAG 2430
 RESULT 15
 ID ADC08208 standard; DNA; 2451 BP.
 XX ADC08208;
 DT 18-DEC-2003 (first entry)
 XX Rice DNA sequence Seq ID513 related to grain filling.
 DE plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarcane;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KW gene; ds; plant.
 XX Oryza sativa.
 OS WO200300905-A2.
 PN 03-JAN-2003.
 PD 21-JUN-2002; 2002WO-1B002450.
 PF 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Kreps J, Provart N, Riche D;
 DR WPI; 2003-229341/22.
 DR P-PSDB; ADC08209.
 PS Claim 35; SEQ ID NO 513; 130pp; English.
 XX
 CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarcane, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more
 CC desirable level using the present invention. The present sequence is a
 CC DNA sequence encoding a rice protein of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/publishedpc_sequences.
 CC
 XX Sequence 2451 BP; 557 A; 674 C; 628 G; 592 T; 0 U; 0 Other;
 SO
 Query Match 49.1%; Score 1287.6; DB 9; Length 2451;
 Best Local Similarity 71.4%; Pred. No. 0;
 Matches 1724; Conservative 0; Mismatches 685; Indels 6; Gaps 2;
 OY 7 GAGGTGCTCTCACTCCGCTCCACAGTCTCCGTGAGGCTTTGATGAGACCTCTTGCT 66
 DB 19 GACCGCGTCTGAGCCGCTCCACAGGCTGAGGAGGCACTCGGCACTCTCTCGCC 78
 OY 67 CACAGAAAGAGATTTGGCTTGCTCTCAAGATCGAGGCAAGAAAGAAATCTG 126
 DB 79 CACCCCAATAGCTCGTGGCGCTTCAACAGGCTGTTAACTCGAAGAAAGAAATCTT 138
 OY 127 CAAACCATCAATATTATTCTAGAGT---TTGAAGCTATCCCTGAAGAAACAGAAAGAG 183
 DB 139 CAGGCCCAAGATCATTTGCTGAGTACAAACGCAATCTGAGGCTGATCGTGAAGAG 198
 OY 184 CTGCTATAGTGTGATTTTGAAGTATGGAAGCTAGTCAAGAGGAGTCTGTCTT 243
 DB 199 CTGAAGACGATGCTTTGAGAGATGCTGAGAGTCAAGAGAGGAGATTTATCTCC 258
 OY 244 CCATGGGTTGCACTGCTGTTGTCGAAGGCTGTTGGAGTACATTAGAGTAAT 303
 DB 259 CATGGGTTGCTCTTGCATTCGCCCAAGGCGGATCTGGAGATATGAGGCTCAAT 318
 OY 304 GTTACGCGCTTGTGTTGAGAACTCATGTTGCTGATCTCACTTCAAGAAAG 363
 DB 319 GTGAGCAAGCTCGTGTGAGTGTGATGCTGACAGTCCCGAGTACTTGAGTTCAAGAAAG 378
 OY 364 CTGTGATGAGAGTTCAATGAAACTTTGTTGAAATTTGAACTTTGAGCCTTCAAC 423
 DB 379 CTGTGAGAAAGAGCAACAAACACTTTGTGCTTGAAGCTTGAAGTTCAGCCATTCAAT 438
 OY 424 TCATCATTTCCCGCCCAAGCTCTTCAAAATCCATTGATGATGAGTTCCTTAAT 483
 DB 439 GCTCTCTTCCCTGCTCTTCTCTGTCGAGTCCATTGAGTGTGAGTCTCTCAAC 498
 OY 484 CGTCACTTTTCGCAAAATTTTCCATGACAGAGAGATGACCTTTGCTGAATTC 543
 DB 499 AGGATGTGTCATGAGAGCTCTTCCATGACAGAGAGATGATCCACAGCTTCACTTC 558
 OY 544 CTGAGTCTCATTTGTCAGAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAACTTG 603
 DB 559 CTTCGTCAACAACTCAAGAGGCTATGATGATGATGATGACAGAAATTCAGAACTTG 618
 OY 604 AATGCTCTCAAGATTTTGAAGAAAGAGAGATGATCTGTACCTTACCTCTGAG 663
 DB 619 AGTGTCTCCAGAGTCTGAGAGAGCTGAGAGCACTTTCTGCTTTTCAGCAGAC 678
 OY 664 AACCATGTCGCAATTTGAAACACCGGTTCCAGAAATCGTTTGAAGAGTTGGGCT 723
 DB 679 ACCCTTACTAGAAATTCACACAGGTTCCAGAACTTGTGTCGAGAGAGGTTGGGCT 738
 OY 724 GACACCGGAGAGCGGCTGAGATATCAACTCTTTTGAATCTTTTGAAGCAACT 783
 DB 739 GACTGTCCAAAGCTTCCAGAGACCAATTCACCTCTTTTGAAGCACTTTTGAAGCCCT 798
 OY 784 GATCTTGAACCCCTTGAAGATTTCTTGGGAGATCCCATGAGTCAATGTTGTGAT 843
 DB 799 GATCTGTCAACCTTCGAGAGTCTTGAAGCAATCCCAATGTTGTTCAATGTTGTCAAT 858
 OY 844 CTCACTCCCAAGAGATCTTCTCAAGACATGTTTGGGATATCCCAACCGATGGG 903
 DB 859 ATGTCCTCAACCGTATCTTCTGAGCAAGATGTTGGGATATCTTGAACAGAGG 918
 OY 904 CAGTTTGTTCATCTTGTGATCAAGTCCGAGCTTTGAGAAATGAGATCTCTCCGATA 963
 DB 919 CAGTTGTTCATCTTGTGATCAAGTCCGAGCTTTGAGAAATGAGATGAGATCTCTCCGATA 978

QY 964 AAGCAACAGGACCTCAACATGACCCCTGGATTCCTATTATTACTAGACTTCTTCGAT 1023
Db 979 AAGGACGAGGCTCTCAATATTAACACGACGATCCTTATGTGCACAGGCTCTCCGAT 1038
QY 1024 GCTGCGGAACAATGCGGTCAACGACTTGAGAAAGTATACGGAACAGACACTCGAT 1083
Db 1039 GCACTGGACACACTGGGTCTAGGCTTTGAGAAAGTCTTGACCGACACACCCAC 1098
QY 1084 ATTCTTGAGTACCTTCAGAACAGAAAGGAAATGTTGCAAAATGATCTCAAGATT 1143
Db 1099 ATCTCAAGTGTGCTTTAGAACTGAAATGGAATGTTCCAAATGATCTCAGTTT 1158
QY 1144 GAAAAAGTGGCCATCTATTTGAAAACTCAACAGAGATGTTGCTCATGAAATCTCCAA 1203
Db 1159 G---AGGCTGGCGGTACTGAAAACATCTAGTGTGTCGGCAAGAAATGCTGGA 1215
QY 1204 GAGTTGACGGGACGCCGATCTGATCATCGAAAAACNACAGCGCAATATGTCGCC 1263
Db 1216 GAGCTCCAGGCCAATCTGACTTGATCATCGAAACTACAGTACGGAACCTTGTTGCA 1275
QY 1264 TCCTTGTGCGACATTAATTAAGTGTCAACAGTGCACACATGCGCCATGCTTTGAGAG 1323
Db 1276 TGCTTGTGTGACCAAGATGGGTGTACCATGTATCATTTGCCCATGGCTTGAAGAA 1335
QY 1324 ACAAAATATCCAGATTCAATATCTATTGGAAGAGCTTGAAGCAATATCATTTCTCT 1383
Db 1336 ACCAAGTACCCCACTGACCTCTACTGAAAGATTTGAGATCACTACACTTCTCA 1395
QY 1384 TGCCAAATTACAGTGATCTTTTTCATGAAACCATACAGATTCATCATCAGCAT 1443
Db 1396 TGCCAGTTCAAACTGACTTGATTTGCAATGAACACGCTGACTCATCATACAGTACC 1455
QY 1444 TTCCAGAAATTTGACAGAAAGACACTGTGTCAATACGAGGCCACACTGCTTC 1503
Db 1456 TTCCAGAGATTTGCGGAAACAGGACACTGTGTGATGATGATCATGAGCATTC 1515
QY 1504 ACTCTTCTGGTCTCTACCGTGTGTATGATGATGATGATGATGATGATGATGATGAT 1563
Db 1516 ACCATGCTGT 1575
QY 1564 ATTGTTCCCTGGT 1623
Db 1576 ATTGTCTGCTGGT 1635
QY 1624 TTGAAGCATTTCACTCTGAGATCGAAGACCTTTTACACAAAGTTGAGATGAAAG 1683
Db 1636 CTGACCTCCCTCAACCCAGAGATTTGAGAGTGTGCTACAGGAAATTGACACACAG 1695
QY 1684 CACTTATGT 1743
Db 1696 CACAAGTTTATGT 1755
QY 1744 CGTGTCAAGAACTTAACCGGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1803
Db 1756 CGTGTCAAGAACTTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1815
QY 1804 TTGCTTAACCTCTGATGT 1863
Db 1816 CTGGTTAACCTTGT 1875
QY 1864 AAGCTGAATGAAGAAATGTTTGAAGTGTGACAAAGTACAACTTGAACGGCCATTC 1923
Db 1876 CAGCTGAGTTCAAGAAATGTTGACCTTATTTGAGCAATACAACTTGAATGGCCACATC 1935
QY 1924 AGATGATATCATCTCAATGAAAGAAATCGAAATGTTGAATTTACGATACATTTGC 1983
Db 1936 CGCTGATCTTCGCGAGATGAAACGTGTCCGCAACGTTAGCTTACCTTTAATCTGC 1995
QY 1984 GACACGAAGGTGCTTTGTACAGCTGATGTATGAAAGCTTTGATGACAGTTGTG 2043
Db 1996 GACACGAAGGTGCTTTGTGACAGCGCTTTCTACAGAGCATTTGATGATGATGATGAT 2055
QY 2044 GAGCAATGACTTGGGTTTGTGCAATTCGCACTGTAAAGGTGACACGCGAGATT 2103

Db 2056 GAGTTCATGACCTGTGTCTCCGACATTCGCAACGCCCTATGTGTGTCCAGCTGAGATC 2115
QY 2104 ATTGTCCATGGAAATCTGTTTCAACATTTGATCTTTACATGATGTGATCAAGCTGTGAC 2163
Db 2116 ATGTTGAACGAGATGTGTGCTTCCACATTTGACCCATACCAAGGTGTCAAGGCTCGGG 2175
QY 2164 ATACTGTGATTTCTTTGAAAAGTGTAAAGATGATCTCACTGGATTAAGATCTCC 2223
Db 2176 CTGCTGTGTAGTCTTTTGAAGTGTGCAAGAGACCCAGCCACTGACCAAGATCTCG 2235
QY 2224 CAAGGAGCTTTGAAGAAATGAGAGAGATATACATGAAAGATTACTCGAGAGACTA 2283
Db 2236 CAGGCGGTCTTACGCTATGAGAGAAATACACTGAAAGCTTACTGTGAGAGGCTG 2295
QY 2284 TTGACCTGACAGAGTGTATGATTTCTGGAAGCATGTTTCAACCTTGAACGCCGTGAG 2343
Db 2296 ATGACCTCAACGCGCTTACAGGATTTCTGAAAGTGTGAACTGTGAAAGGCTGAA 2355
QY 2344 AGTGTGTTACCTTGAGATGTTTATGTCTTAAAGTACCGTAAAGCTGCTGAATCACTT 2403
Db 2356 ACCGCGCTTACCTTGAGATGCTGTATGCCCTCAAGTACGCAACATGGCTAGCACCGTC 2415
QY 2404 CCAATGSCAAGAG 2418
Db 2416 CCGTGGCGGTGAG 2430

Search completed: June 5, 2004, 01:13:49
Job time : 1477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:48:19 / Search time 51 Seconds

(without alignments)
4986.425 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247

Sequence: 1 MAERALTRVHSLRRLDETL.....EMFYALKYRKLAESVPLAEE 806

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Listing first 45 summaries

SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_todent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virins.*
16: sp_bacterioph.*
17: sp_archaeop.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	10 Q9XGB7	Q9XGB7 gossypium h
2	3839.5	90.4	805	10 Q9SLR2	Q9SLR2 citrus unsh
3	3815.5	89.8	805	10 Q9SLY1	Q9SLY1 citrus unsh
4	3674.5	86.5	806	10 Q9TOM9	Q9TOM9 pisum sativ
5	3673.5	86.5	806	10 Q9SBL8	Q9SBL8 citrullus l
6	3662.5	86.2	806	10 Q9SLR3	Q9SLR3 phaseolus v
7	3649.5	85.9	805	10 Q9TOM6	Q9TOM6 medicago tr
8	3637.5	85.6	805	10 Q9XG65	Q9XG65 medicago tr
9	3614.5	85.1	808	10 Q9LXL5	Q9LXL5 arabidopsis
10	3536.5	83.3	804	10 Q9AYR8	Q9AYR8 pisum sativ
11	3496.5	82.3	805	10 Q82633	Q82633 lycopersico
12	3488.5	82.1	805	10 Q7Y078	Q7Y078 solanum tub
13	3470.5	81.7	805	10 Q84UC3	Q84UC3 solanum tub
14	3452	81.3	803	10 Q9LMB7	Q9LMB7 chenopodium
15	3443.5	81.1	805	10 Q82691	Q82691 lycopersico

17	3296	77.6	816	10 Q8W1W4	Q8W1W4 bambusa old
18	3293.5	77.5	816	10 Q8LJ74	Q8LJ74 x mokara cv
19	3292	77.5	816	10 Q43706	Q43706 zea mays (m
20	3289.5	77.5	808	10 Q8W1W3	Q8W1W3 bambusa old
21	3287.5	77.4	808	10 Q8W1W2	Q8W1W2 bambusa old
22	3272	77.0	816	10 Q8G523	Q8G523 oryza sativ
23	3270.5	77.0	816	10 Q8LJ75	Q8LJ75 oncidium cv
24	3264.5	76.9	802	10 Q9LKR0	Q9LKR0 sacccharum o
25	3209.5	75.6	815	10 Q43223	Q43223 trifolium ae
26	3186.5	75.0	808	10 Q82073	Q82073 trifolium ae
27	3071	72.3	811	10 Q9SLY2	Q9SLY2 citricum unsh
28	3056	72.0	811	10 Q9SLV8	Q9SLV8 citrus unsh
29	2996	70.5	811	10 Q84T18	Q84T18 solanum tub
30	2972	70.0	809	10 Q8L5H0	Q8L5H0 zea mays (m
31	2962	69.7	809	10 Q9M111	Q9M111 arabidopsis
32	2956	69.6	796	10 Q93W53	Q93W53 zea mays (m
33	2953	69.5	822	10 Q94G60	Q94G60 beta vulgar
34	2940	69.2	812	10 Q9FRX3	Q9FRX3 pyrus pyrif
35	2934	69.1	811	10 Q9ZPC5	Q9ZPC5 crataegus
36	2892	68.1	808	10 Q9SBD5	Q9SBD5 arabidopsis
37	2837	66.8	809	10 Q9ZPC6	Q9ZPC6 crataegus
38	2487.5	58.6	532	10 Q94CC8	Q94CC8 arabidopsis
39	2312.5	54.5	942	10 Q9FXJ2	Q9FXJ2 arabidopsis
40	2188.5	51.5	454	10 Q9ZRC4	Q9ZRC4 gossypium h
41	2150.5	50.6	798	10 Q7XXN5	Q7XXN5 oryza sativ
42	2148.5	50.6	798	10 Q7XXN1	Q7XXN1 oryza sativ
43	2148	50.6	887	10 Q9FHU4	Q9FHU4 arabidopsis
44	2090.5	49.2	794	10 Q820M5	Q820M5 nicotiana t
45	1821.5	42.9	422	10 Q8W402	Q8W402 nicotiana t

ALIGNMENTS

RESULT 1
ID Q9XGB7 PRELIMINARY: PRT: 806 AA.

AC Q9XGB7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Sucrose synthase.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucotyledons II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Acacia SJ-2; TISSUE=Fiber;
RA Perez-Grau L., Delmer D.;
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U73588; AAD28641.1; -;
DR GO: GO:0005058; P:biosynthesis; IEA.
DR GO: GO:0005985; P:sucrose metabolism; IEA.
DR InterPro: IPR001296; Glyco_transf.1.
DR InterPro: IPR000368; Sucrose synth.
DR Pfam: PF00862; Sucrose synth.1.
SQ SEQUENCE 806 AA; 92754 MW; 5916FA6B7E0DABEC CRC64;

Query Match 100.0%; Score 4245; DB 10; Length 806;
Best Local Similarity 100.0%; Pred. No. 1.3e-302;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAERALTRVHSLRRLDETLAHRNELLALSRLEGKGGILOHQTLEPEALPEERK 60
DB 1 MAERALTRVHSLRRLDETLAHRNELLALSRLEGKGGILOHQTLEPEALPEERK 60
QY 61 KLANGAFPEVVKASQEAIVLPWPVALAVRPBGWMEYIRVVNVALVVEELTVAEYLHPE 120
DB 61 KLANGAFPEVVKASQEAIVLPWPVALAVRPBGWMEYIRVVNVALVVEELTVAEYLHPE 120

```
QY 121 ELVDGSSNGNFVLELDEFPENSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
DB 121 ELVDGSSNGNFVLELDEFPENSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
QY 181 FLRVCHKGKMMMLNDRIQNLALOHVLRKAEEYIGTLPRPTPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHKGKMMMLNDRIQNLALOHVLRKAEEYIGTLPRPTPCAEFEHRFOEIGLERGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLGYPTDG 300
DB 241 GDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLGYPTDG 300
QY 301 GQVYIILDQVRLLENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS 360
DB 301 GQVYIILDQVRLLENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS 360
QY 361 DILRVPRTEKGIVRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNXSDGNIV 420
DB 361 DILRVPRTEKGIVRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNXSDGNIV 420
QY 421 ASLNAHKLGVTOCTIAHALEKTKYPSDIIYWKLEDKYHFSQFTADLFAMNHTDFITTS 480
DB 421 ASLNAHKLGVTOCTIAHALEKTKYPSDIIYWKLEDKYHFSQFTADLFAMNHTDFITTS 480
QY 481 TFOEIASGKDTVGQYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
DB 481 TFOEIASGKDTVGQYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
QY 541 RLKHFHPEIBDLTYKYENEBELCYLNDNRNPILEFTMRDLRVKNLTGLVEMCGNPLXR 600
DB 541 RLKHFHPEIBDLTYKYENEBELCYLNDNRNPILEFTMRDLRVKNLTGLVEMCGNPLXR 600
QY 601 ELANLVVVGGRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNRIRINVELYRYI 660
DB 601 ELANLVVVGGRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNRIRINVELYRYI 660
QY 661 CDTKGAFOVOPALYEAFGITVVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
DB 661 CDTKGAFOVOPALYEAFGITVVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTKWIKISQRLITLTGVYGFPMKHSNLER 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTKWIKISQRLITLTGVYGFPMKHSNLER 780
QY 781 ESRRYLEMFVALKYRKLAEVPLAE 806
DB 781 ESRRYLEMFVALKYRKLAEVPLAE 806

RESULT 2
Q9SL52 PRELIMINARY; PRT; 805 AA.
ID Q9SL52
AC Q9SL52
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Sucrose synthase.
GN CITSUS1-2.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Miyagawa-wase;
RA Komatsu A.;
RT "Cloning of sucrose synthase (CITSUS1) gene from Citrus.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR GO; GO:0009058; P-biosynthesis; IEA.
DR GO; GO:0005985; P-sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
```

```
DR InterPro; IPR000368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose_synth; 1.
SQ SEQUENCE 805 AA; 92172 MM; 2098FC1F524C266E CRC64;

Query Match 90.4%; Score 3839.5; DB 10; Length 805;
Best Local Similarity 90.7%; Pred. No. 7.2e-273;
Matches 731; Conservative 34; Mismatches 40; Indels 1; Gaps 1;

QY 1 MAERALRVISLRRLDETTLAHREITLALSRLEGKGLIOHIOHILEFEALPEERK 60
DB 1 MAERALRVISLRRLDETTLAHREITLALSRLEGKGLIOHIOHILEFEALPEERK 60
QY 61 KLANAFEEVLKASQEAIVLPVVALAVRPRGVMEYIRVNVHVLVEELTVAEYLHKE 120
DB 61 HLTGAFGEVLRAQEAIVLPVVALAVRPRGVMEYIRVNVHVLVEELTVAEYLHKE 120
QY 121 ELVDGSSNGNFVLELDEFPENSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
DB 121 ELVDGSSNGNFVLELDEFPENSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
QY 181 FLRVCHKGKMMMLNDRIQNLALOHVLRKAEEYIGTLPRPTPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHKGKMMMLNDRIQNLALOHVLRKAEEYIGTLPRPTPCAEFEHRFOEIGLERGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLGYPTDG 300
DB 241 GDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLGYPTDG 300
QY 301 GQVYIILDQVRLLENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS 360
DB 301 GQVYIILDQVRLLENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS 360
QY 361 DILRVPRTEKGIVRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNXSDGNIV 420
DB 361 DILRVPRTEKGIVRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNXSDGNIV 420
QY 421 ASLNAHKLGVTOCTIAHALEKTKYPSDIIYWKLEDKYHFSQFTADLFAMNHTDFITTS 480
DB 421 ASLNAHKLGVTOCTIAHALEKTKYPSDIIYWKLEDKYHFSQFTADLFAMNHTDFITTS 480
QY 481 TFOEIASGKDTVGQYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
DB 481 TFOEIASGKDTVGQYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
QY 541 RLKHFHPEIBDLTYKYENEBELCYLNDNRNPILEFTMRDLRVKNLTGLVEMCGNPLXR 600
DB 541 RLKHFHPEIBDLTYKYENEBELCYLNDNRNPILEFTMRDLRVKNLTGLVEMCGNPLXR 600
QY 601 ELANLVVVGGRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNRIRINVELYRYI 660
DB 601 ELANLVVVGGRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNRIRINVELYRYI 660
QY 660 CDTKGAFOVOPALYEAFGITVVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 719
DB 660 CDTKGAFOVOPALYEAFGITVVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 719
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTKWIKISQRLITLTGVYGFPMKHSNLER 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTKWIKISQRLITLTGVYGFPMKHSNLER 780
QY 781 ESRRYLEMFVALKYRKLAEVPLAE 806
DB 781 ESRRYLEMFVALKYRKLAEVPLAE 806

RESULT 3
Q9SLY1 PRELIMINARY; PRT; 805 AA.
ID Q9SLY1
AC Q9SLY1
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
```

DE Sucrose synthase.
GN CITSUS1.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids I; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cy. Miyagawa-Wase; TISSUE=fruit;
RA Komatsu A.;
RT "Isolation of three sucrose synthase isoforms from Citrus."
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB022092; BAA89905.1; -
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose_synth_1.
SO SEQUENCE 805 AA; 9211 MW; E2204B24E608328 CRC64;

Query Match 89.8%; Score 3815.5; DB 10; Length 805;
Best Local Similarity 90.3%; Pred. No. 4,2e-271;
Matches 728; Conservative 35; Mismatches 42; Indels 1; Gaps 1;

QY 1 MAERALTIVHSIRERLDTLLAHRNEIILALSRIGKKGKGILOHQIILFEPAIPEENRK 60
DB 1 MAERALTIVHSIRERLDTLLAHRNEIILALSRIGKKGKGILOHQIILFEPAIPEENRK 60
QY 61 KLANGAFEEVLKASQEAIVLPWPVALAVPRPGWEXIRVNVHVALVVELTVAEYLAFKE 120
DB 61 HLTGAFGEVLRATGEAIVLAPWVALAVPRPGWEXIRVNVHVALVVELTVAEYLAFKE 120
QY 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
DB 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
QY 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
DB 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
QY 241 GDTAERVLEMIOQLLDLEATDPTCLEKFLGRIPWVFNVLITPHGYFAODNVLAGYPTDG 300
DB 241 GDTAERVLEMIOQLLDLEATDPTCLEKFLGRIPWVFNVLITPHGYFAODNVLAGYPTDG 300
QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGQRLKEXYGTSHS 360
DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGQRLKEXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420
DB 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420
QY 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420
DB 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420
QY 421 ASLLAHKLGVTQCTAAHLEKTKYPSDSIYWKMLDEKDKHFSQGFADIPANNHDTFIIITS 480
DB 421 ASLLAHKLGVTQCTAAHLEKTKYPSDSIYWKMLDEKDKHFSQGFADIPANNHDTFIIITS 480
QY 481 TFOEIASGKDVYGOYESHATAFTLPGLYRVVHGIDVFDPKFNIVSPGADMEIYFPYTEEKR 540
DB 481 TFOEIASGKDVYGOYESHATAFTLPGLYRVVHGIDVFDPKFNIVSPGADMEIYFPYTEEKR 540
QY 540 TFOEIASGKDVYGOYESHATAFTLPGLYRVVHGIDVFDPKFNIVSPGADMEIYFPYTEEKR 540
DB 540 TFOEIASGKDVYGOYESHATAFTLPGLYRVVHGIDVFDPKFNIVSPGADMEIYFPYTEEKR 540
QY 541 RLKHFHPEIEDLLTKVENEBHLVCLDRNKPILFTFPRDLRVKNLTGLVWCGKNPKLR 600
DB 541 RLKHFHPEIEDLLTKVENEBHLVCLDRNKPILFTFPRDLRVKNLTGLVWCGKNPKLR 600
QY 600 ELVNVLVVGGDRKESKDLSEKAEKMKPELIDKXNLNGOFRTSSQNNRIIRNVELYVI 660
DB 600 ELVNVLVVGGDRKESKDLSEKAEKMKPELIDKXNLNGOFRTSSQNNRIIRNVELYVI 660
QY 661 CDTKGAFYOPALYEAFFGLTVVEAMTCGLPTFATCGGPAEIIIVHSGSFNIDPYHGDQA 720
DB 661 CDTKGAFYOPALYEAFFGLTVVEAMTCGLPTFATCGGPAEIIIVHSGSFNIDPYHGDQA 720
QY 660 CETKGAFYOPALYEAFFGLTVVEAMTCGLPTFATCGGPAEIIIVHSGSFNIDPYHGDQA 719

QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEEXYTWKYEERLITLGYGFWKVSNLERR 780
DB 720 EILDVDFEKKCKDPSHMDKISQGLKRIEEXYTWKYEERLITLGYGFWKVSNLERR 779
QY 781 ESRRLYEMFYALKYRKLAESVPLAE 806
DB 780 ESRRLYEMFYALKYRKLAESVPLAE 805

RESULT 4
Q9TOM9 PRELIMINARY; PRT; 806 AA.
ID Q9TOM9
AC Q9TOM9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Sucrose synthase (EC 2.4.1.14).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RA Craig J., Barratt P., Tatge H., Dejardin A., Handley L., Gardner C.D.,
RA Barber L., Wang T., Hedley C., Martin C., Smith A.M.;
RT "Mutations at the rug4 locus alter the carbon and nitrogen metabolism
of pea plants through an effect on sucrose synthase";
RL Plant J. 17:353-362(1999).
DR EMBL; AJ012080; CA09910.1; -
DR GO; GO:0046524; F:sucrose-phosphate synthase activity; IEA.
DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose_synth_1; 1.
DR Glycoalyt;transf;transferase.
SO SEQUENCE 806 AA; 92501 MW; CD20A7878C46F768 CRC64;

Query Match 86.5%; Score 3674.5; DB 10; Length 806;
Best Local Similarity 86.4%; Pred. No. 9.3e-261;
Matches 696; Conservative 52; Mismatches 57; Indels 1; Gaps 1;

QY 1 MAERALTIVHSIRERLDTLLAHRNEIILALSRIGKKGKGILOHQIILFEPAIPEENRK 60
DB 1 MATBRLTRVHSIRERLDTLLAHRNEIILALSRIGKKGKGILOHQIILFEPAIPEENRK 60
QY 61 KLANGAFEEVLKASQEAIVLPWPVALAVPRPGWEXIRVNVHVALVVELTVAEYLAFKE 120
DB 61 KLTGAFGEVLRATGEAIVLAPWVALAVPRPGWEXIRVNVHVALVVELTVAEYLAFKE 120
QY 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
DB 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
QY 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
DB 121 ELVDGSSNGNVLLEDFEPFNSSPPRPTLSKISGNGVEFLNRHISAKLFHOKESHAPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLENGW 240
QY 241 GDTAERVLEMIOQLLDLEATDPTCLEKFLGRIPWVFNVLITPHGYFAODNVLAGYPTDG 300
DB 241 GDTAERVLEMIOQLLDLEATDPTCLEKFLGRIPWVFNVLITPHGYFAODNVLAGYPTDG 300
QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGQRLKEXYGTSHS 360
DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGQRLKEXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420
DB 361 DILRVPTTEKGIYRKWISREKWPYLETYTEDEVAHEISKELHGTPLDITGNXSDGNIY 420

D	b	361	HIIRVFRPODKIVRKMIISRF-VMPELETYEDVAHAEALOGKEDLIVGNISDNIY	4129
Q	y	421	ASLHAKLGVOTCTIAHALEKTRYPDSDIYWKQJEDKXHFSCQFTADLFAMNHTDFIITS	480
D	b	420	ASLHAKLGVOTCTIAHALEKTRYPDSDIYWKQJEDKXHFSCQFTADLFAMNHTDFIITS	479
Q	y	481	TFQELAGSKOTVGOYESHTAFTLPGLYRVVHGIDVDFKFNIVSPGADMEIYPPYTEKR	540
D	b	480	TFQELAGSKOTVGOYESHTAFTLPGLYRVVHGIDVDFKFNIVSPGADQTIYFPYETSR	539
Q	y	541	RLKHFEIIEDLLIVTVEBEHLCVLIDNRKCIILFTMRPLDRVKNLTGLVWCSKQEKLR	600
D	b	540	RLTSPFEIIEILLIVTVEBEHLCVLIDNRKCIILFTMRPLDRVKNLTGLVWGYSKNAKLR	599
Q	y	601	ELANLVVVGSDRRKESKDLBEKAEKMKQFELLIDKTNLNGQRMVLISSQNMRRINVELYRI	660
D	b	600	ELVNLVVVGSDRRKESKDLBEKAEKMKQFELLIDKTNLNGQRMVLISSQNMRRINVELYRI	659
Q	y	661	CDPKGAFVOPALYEAAGLTVVEAMTCGLPTFATCNGSPALIIYHGKSGFNIDPRHGOAA	720
D	b	660	CDPKGAFVOPALYEAAGLTVVEAMTCGLPTFATLNGSPALIIYHGKSGFNIDPRHGOAA	719
Q	y	721	DIIVDFPEKKQDPSMDKISOGGLRIEEXKTMKIYSEBRLTLTGUVGFWMKHSNLER	780
D	b	720	DLIVDFPEKKQDPSMDKISOGGLRIEEXKTMKIYSEBRLTLTGUVGFWMKHSNLDRL	779
Q	y	781	ESRRYLEMFALKYRKLAAESVPLAE	806
D	b	780	ESRRYLEMFALKYRKLAAESVPLAE	805

RESULT 5

ID	Q9SBL8	PRELIMINARY;	PRT;	806 AA
----	--------	--------------	------	--------

AC Q3386; 01-May-2000 (TREMBLrel. 13, Created)
 DT 01-May-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-May-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Mus.
 GN MSUS.
 OS Citrullus lanatus (Watermelon) (Citrullus vulgaris).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.
 NCBI_TaxID=3654;

RA SEQUENCE FROM N.A.
 RP Kim J., Kang H., Jun S., Lee J., An G.,
 RT "watermelon mRNA for Gs3-regulated in developing seeds, complete cds
 RT (Cloning of three gibberellin-regulated cDNAs from watermelon during
 RT early seed development: down-regulated one cDNA and up-regulated two
 RT cDNAs).";
 RT Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL; AB015561; EA88532.1; 1.
DR GO; GO:0003055; P:ribosynthesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose synth.
DR Pfam; PF00534; Glycosyltransferase_1; 1.
DR Pfam; PF00862; Sucrose synth. 1.
DR SEQUENCE; 806 AA; 9255 MW; 4FDBECAD6C140337 CRC64;

Query Match 86.5%; Score 3673.5; DB 10; Length 806;
Best Local Similarity 86.8%; Ped. No. 1.1e-260;
Matches 698; Conservative 48; Mismatches 57; Indels 1; Gaps 1

Qy	1	MAEALTRVNSIAERIDETILMARNTIALLSRIEGKGGIIONHQTLLSEALPEENRK	60
Db	1	MAEVLNRVSHLERIDETILIAQRNITLIGLSIKTEKGIGIIONHQLAEALPEENRK	60
Qy	61	KLANGAFEEYLKASQEAIVLPVVALAVRPRGWEYIRVNVHVALVEELTVAAYLHPKE	120
Db	61	KLADGDGEVLRVTOESIVLPVVALAVRPRGWEYIRKONVHVALVEELOVAAYLHPKE	120

QY	121	LVNDSSNGNVTLEDEEPNUSFERPTLSKISGNGVEINHLASKLPHDKSMHPLLE	180
Db	121	ELVNDSSNGNVTLEDEEPFNASFPRLTUSKISGNGVEINHLAKLPHGKSMOPLLD	180
QY	181	FLRVHCHGKNMMLNDR1QNLNALQHVLRKAEYLGTLPEETPCAEFERFOEIGDERGW	240
Db	181	FLRVHCHGKNMMLNDR1QNLNAFQHVLRKAREYLATLAPETIPYSEFANKFEIIGDERGW	240
QY	241	GDYARVLEMIQLDLDDLEATDPCLTLEKPLGSIIPWFENVVILTPHGYFADONVLTGPDTC	300
Db	241	GNTARVLEMIQLDLDDLEADPCCTFEKPLGSIIPWFENVVILSPHGYFADONVLYGPDTC	300
QY	301	GOVVYIILDOVALENEMILRIKQOGLNIPRLLITRLPLDAVGTCCGGRLEKYTETHS	360
Db	301	GOVVYIILDOVALALEHMIORIKQOGLDITRPLILITRLPLDAVGTCTQGRLEKFCGTEHS	360
QY	361	DILRVPTTEKGIYKRWISRFKWPYLETTEDVAHEISKELHGTPLDLIGNXSGDNIV	420
Db	361	HILRVPTNEKGIYKRWISRPE-VWPYLETYTEBVAQOELTKELOGRPDLIGNYSGDNIV	419
QY	421	ASLAKHKGVTQCTTAHLAEKTKYPSDIPYMKLEBKHYFSCQFTADLFAMNHTDPLITS	480
Db	420	ASLAKHKGVTQCTTAHLAEKTKYPSDIPYMKRPDDKHYFSSQFTADLFAMNHTDPLITS	479
QY	481	TFQEIAGSKDITVGQYESHSTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKR	540
Db	480	TFQEIAGSKDITVGQYESHSTAFTLPGLYRVVHGIDVDPFNITSPGADMSIYFPYTEK	539
QY	541	RLKHHPREIEDLTYRVKNEEHLCVLINDNKILFPMPLIDRYONLTGVEMGOKPKLR	600
Db	540	RLTSHPEIEELLYSEVENEEHLCVLKDORSKILFPMALIDRYONLTGLVEMYGKXKRLR	599
QY	601	ELANLVVVGDRKESKDLBEKAEMKKMFELIDKNLNGQFHWISQOMNRIRNVELYRI	660
Db	600	ELVNLVVVAGDRKESKDNBEKAEMEBKYLILKTYLNGQFRWISQOMNRVRNGEYRCI	659
QY	661	CDTKAPFOPALYKAPGLTVYBAMTCGLPTATCNGBPALIIYHSGSGFIDIPYHGOAA	720
Db	660	ADTKAPFOPAYAYKAPGLTVYBAMTCGLPTATCNGBPALIIIDGSGSGFIDIPYHGRAA	719
QY	721	DILVDFEKKCKDPAMWDKISOGGAKRIIEKXTWKIYSERLLTLTGCVGFMKIVSNLER	780
Db	720	EILVDFEKKSKDPAMWDKISOGGAKRIIEKXTWQIYSERLLTLTGCVGFMKIVSNLRL	779
QY	781	ESRRYLEMFYALKYKRLAESVPLA 804	
Db	780	ESRRYLEMFYALKYKRLADSVPOA 803	

RESULT 6

ID	081610	PRELIMINARY;	PRT;	806	AA
----	--------	--------------	------	-----	----

AC *Osar10*; 01-NOV-1998 (TREMblrel). 08, Created)
 DT 01-NOV-1998 (TREMblrel). 08, Last sequence update)
 DT 01-JUN-2003 (TREMblrel). 24, Last annotation update)
 DE Module-enhanced sucrose synthase (EC 2.4.1.13).
 GN NESS.
 OS *Platanus sativum* (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 OC eustrods I; Fabales; Fabaceae; Papilionoideae; Viciae; Platan.
 OC NCBI TaxID=3888;

RP SEQUENCE FROM N.A

RA Fedorova M., Tikhonovich I.A., Vance C.P.

RT nodules. I. In situ localization in the effective nodules.";

DR EMBL; AF079851; AAC28107.1; -.

DR GO:0016757; F:transferrase activity, transferring glycosyl. . . ; IEA;
DR GO:0009058; P:biosynthesis; IEA.
DR GO:0005985; P:sucrose metabolism; IEA.
DR InterPro: IPR001296; Glyco_trans_1.
DR InterPro: IPR00368; Sucrose_synth.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Pfam: PF00862; Sucrose_synth_1.
DR Glycoyltransferase; Transferrase.
SQ SEQUENCE 806 AA; 92413 MW; 7B9B52F3F0C60D1 CRC64;

Query Match 86.2%; Score 3662.5; DB 10; Length 806;
Best Local Similarity 86.2%; Pred. No. 7.1e-260;
Matches 695; Conservative 50; Mismatches 60; Indels 1; Gaps 1;

QY 1 MAERALT RVHSRLRRLDETLAHRNEITALLSRLEGKGLIOHQIILFEPAIPENRK 60
DB 1 MATRLTRVHSRLRRLDETLAHRNEITALLSRLEAKGKGLIOHQVIAEFEEIPENRQ 60

QY 61 KLANGAFEEVVKASOEAVLPWPVALAVRPBGWEYLRVNVHVLVVEELVAEYLHKE 120
DB 61 KLIDGAFEEVLRSTOEAVLPWPVALAVRPBGWEYLRVNVHVLVVEELVAEYLHKE 120

QY 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180
DB 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180

QY 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180
DB 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180

QY 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240

QY 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240

QY 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300
DB 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300

QY 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300
DB 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300

QY 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360
DB 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360

QY 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360
DB 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360

QY 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420

QY 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420

QY 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480
DB 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480

QY 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480
DB 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480

QY 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540

QY 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540

QY 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600

QY 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600

QY 601 ELAVLVVVGGRBRKESKLEEKAEKMKFELIDKYNLNGOPRMTSSQNNRIRANVELYXI 660
DB 601 ELAVLVVVGGRBRKESKLEEKAEKMKFELIDKYNLNGOPRMTSSQNNRIRANVELYXI 660

QY 601 ELAVLVVVGGRBRKESKLEEKAEKMKFELIDKYNLNGOPRMTSSQNNRIRANVELYXI 660
DB 601 ELAVLVVVGGRBRKESKLEEKAEKMKFELIDKYNLNGOPRMTSSQNNRIRANVELYXI 660

QY 661 CDTGAFQOPALYEAFGLTVVEAMTCGLPTPATCGNGPAEIIIVHGKSGFNIDPYGDOAA 720
DB 661 CDTGAFQOPALYEAFGLTVVEAMTCGLPTPATCGNGPAEIIIVHGKSGFNIDPYGDOAA 720

QY 661 CDTGAFQOPALYEAFGLTVVEAMTCGLPTPATCGNGPAEIIIVHGKSGFNIDPYGDOAA 720
DB 661 CDTGAFQOPALYEAFGLTVVEAMTCGLPTPATCGNGPAEIIIVHGKSGFNIDPYGDOAA 720

QY 721 DILVDFEKKKDDPSHMDKISQGLKRIEKKYTKIYSERLITLTVGYGFWGHNLEER 780
DB 721 DILVDFEKKKDDPSHMDKISQGLKRIEKKYTKIYSERLITLTVGYGFWGHNLEER 780

QY 721 DILVDFEKKKDDPSHMDKISQGLKRIEKKYTKIYSERLITLTVGYGFWGHNLEER 780
DB 721 DILVDFEKKKDDPSHMDKISQGLKRIEKKYTKIYSERLITLTVGYGFWGHNLEER 780

QY 781 ESRRYLEMFYALKYKLAESVPLAE 806
DB 781 ESRRYLEMFYALKYKLAESVPLAE 806

QY 781 ESRRYLEMFYALKYKLAESVPLAE 806
DB 781 ESRRYLEMFYALKYKLAESVPLAE 806

RESULT 7
Q8GTA3

ID Q8GTA3 PRELIMINARY; PRT; 805 AA.
AC Q8GTA3;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RA Camas A., Silvestre S., Lara M.;
RT "Expression of the gene for sucrose synthase during development of
RT common bean nodules."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF315375; AAN76498.1; -.
DR GO:0016757; F:sucrose synthase activity; IEA.
DR GO:0016757; F:transferrase activity, transferring glycosyl. . . ; IEA.
DR GO:0009058; P:biosynthesis; IEA.
DR GO:0005985; P:sucrose metabolism; IEA.
DR InterPro: IPR001296; Glyco_trans_1.
DR InterPro: IPR00368; Sucrose_synth.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Pfam: PF00862; Sucrose_synth_1.
DR Transferrase; Glycosyltransferase.
SQ SEQUENCE 805 AA; 92024 MW; 94227641C64FE863 CRC64;

Query Match 85.9%; Score 3649.5; DB 10; Length 805;
Best Local Similarity 86.0%; Pred. No. 6.4e-259;
Matches 693; Conservative 48; Mismatches 64; Indels 1; Gaps 1;

QY 1 MAERALT RVHSRLRRLDETLAHRNEITALLSRLEGKGLIOHQIILFEPAIPENRK 60
DB 1 MATRLTRVHSRLRRLDETLAHRNEITALLSRLEAGTGILQHQVIAEFEEIPESRQ 60

QY 61 KLANGAFEEVVKASOEAVLPWPVALAVRPBGWEYLRVNVHVLVVEELVAEYLHKE 120
DB 61 KLIDGAFEEVLRSTOEAVLPWPVALAVRPBGWEYLRVNVHVLVVEELVAEYLHKE 120

QY 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180
DB 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180

QY 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180
DB 121 ELVDSGSGNFVLELDEPPFNSPPTLSKISGVGEFLNRHSAKLFDHKSHPLE 180

QY 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240

QY 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVHCHGKMMILNDRIQNLALOHVLRKAEYIGLTPETPCAEFEHRFOEIGLERGW 240

QY 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300
DB 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300

QY 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300
DB 241 GDTAEVLEMIQLLDLEADPCTLEKFLRIPVFNVLTPHGYAQNVLGYPTG 300

QY 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360
DB 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360

QY 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360
DB 301 GOVVYILDOVALNEMILRIKQOGLNTPRILITRLLPAVGTGCGRLKRYGTEHS 360

QY 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420

QY 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREKWPVLETYEDVAHESKELHGTPLDIGNSDGNIV 420

QY 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480
DB 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480

QY 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480
DB 421 ASLAAHKLGVTOCTTAHLEKTKYPSDIYWKLEBDKTHSCQFTADLFANMHTDFITTS 480

QY 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540

QY 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGYESHTAFTLPGLYRVVGVGIDVPDKFNIVSPGADMEIYFPYTEKR 540

QY 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600

QY 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 600

QY 540 RLTSFHPHEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 599
DB 540 RLTSFHPHEIDLLTYKVENEHLCVLNDNRKPIIFTMRLDRVKNLNGLVEMCGKNPKLR 599

QY 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKTNNGOFRWISSQMNRIYNVELYRYI 660
 DB 600 ELVNLVVVAGDRRKESKDLSEKAEKMKVSLIETTYKLNQCFRWISSQMNRIYNVELYRYI 659
 QY 661 CDTGAFVQPALYEAFLGTVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 720
 DB 660 SDTGAFVQPALYEAFLGTVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 719
 QY 721 DLVDFPEKCKDSDHMDKISQGLKRIEKKYTWKIXSERLLTTLTGYYGFMKAVSNLERR 780
 DB 720 DLVDFPEKCKVESHMDTISQAGLQRIEKKYTWKIXSERLLTTLTGYYGFMKAVSNLERR 779
 QY 781 ESRRYLEMFYALKYRKLAESVPLAE 806
 DB 780 ESRRYLEMFYALKYRKLAESVPLAE 805

RESULT 8

Q9TOM6 PRELIMINARY; PRT; 805 AA.
 AC Q9TOM6: 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Sucose synthase (EC 2.4.1.13).
 GN SUCS1.
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Jemalong; TISSUE=Root nodules;
 RX MEDLINE=99254811; PubMed=10323232;
 RA Hohnjec N., Becker J.D., Pehler A., Perllick A.M., Kuester H.,
 RT "Genomic organisation and expression properties of the MtsucS1 gene,
 RT which encodes a nodule-enhanced sucrose synthase from the model legume
 RT Medicago truncatula.";
 RL Mol. Gen. Genet. 261:514-522 (1999).
 DR EMBL; AJ131943; CAB40794.1; -.
 DR GO; GO:0016157; F:sucrose synthase activity; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR00368; Sucrose_synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucrose_synth; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 805 AA; 92293 MW; 6F1DE955EC3DB1A6 CRC64;

Query Match 85.9%; Score 3646.5; DB 10; Length 805;
 Best Local Similarity 85.6%; Pred. No. 1,le-256;
 Matches 690; Conservative 53; Mismatches 62; Indels 1; Gaps 1;

QY 1 MAERALTIVHSILRELDLTLAHRNEIALLSRLEGKGILOHQLLFEALPEERK 60
 DB 1 MATERLITRVHSILRELDLTLAHRNEIALLSRLEGKGILOHQLLFEALPEERK 60
 QY 61 KLANGAFPEVLKASGEAIVLPVVALAVPRPGWEXIRVNVHALLVEELTVAEYLHKE 120
 DB 61 KLTGGAPEVLRSTGEAIVLPVVALAVPRPGWEXIRVNVHALLVEELTVAEYLHKE 120
 QY 121 ELVNGSSNGNRYLELDFEPNNSPPRPTLSKISNGVFLRHHISAKLPHKESHPLE 180
 DB 121 ELVNGSSNGNRYLELDFEPNNSPPRPTLSKISNGVFLRHHISAKLPHKESHPLE 180
 QY 181 FLRVHCHGKMMANDRIQNTALNOHVARKABEYIGTLPPETPCAEPFHRFOEIGLEKGM 240
 DB 181 FLRLHSYKGTLMANDRIQNPDSLOHVRKABEYISTIDPETPYSEFHRFOEIGLEKGM 240

QY 241 GDTAERVLEMIQALLDLLEATDPCTLEKFLGRIPMVNVNVIITPHGYFAODNVLGYPDTG 300
 DB 241 GDTAERVLESIQALLDLLEAPDPCCTLETFIDRIIPMVNVNVIITPHGYFAODNVLGYPDTG 300
 QY 301 GQVYVILDOVRALENEMLLRIKQGNITPRILITLPLDAVGTGQRIEKKYGTGHS 360
 DB 301 GQVYVILDOVRALSESLRSIKQGLDIRILITLPLPAVGTGQRIEKKYGTGHS 360
 QY 361 DLIRVPRTEKGIYRKWISRFPEKWPYLETTYTBDAHISEKELHGTPLIIGNSDGNIV 420
 DB 361 HILRVPRDRIKGIYRKWISRFPEKWPYLETTYTBDAHISEKELHGTPLIIGNSDGNIV 419
 QY 421 ASLIAHKLGYOCTIAHALKTKYPSDIIYWKLEDKYHSPCOPTADLFAMNHTDFIITS 480
 DB 420 ASLIAHKLGYOCTIAHALKTKYPSDIIYWKLEDKYHSPCOPTADLFAMNHTDFIITS 479
 QY 481 TPQEIAGSKDVGQYESHTAFTLPGLRVNVHGDVPPKFNIVSPGADMEIYFPYTEKR 540
 DB 480 TPQEIAGSKDVGQYESHTAFTLPGLRVNVHGDVPPKFNIVSPGADMEIYFPYTEKR 539
 QY 541 RLKHFHPEIEDLYTKVNEEHLCVLDRNKKPILFTMPRLDRVNLTLGVEMCGKNPLR 600
 DB 540 RLTSFYFIEIELLYSSVNEEHLICVLDRNKKPILFTMPRLDRVNLTLGVEMCGKNPLR 599
 QY 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKTNNGOFRWISSQMNRIYNVELYRYI 660
 DB 600 ELVNLVVVAGDRRKESKDLSEKAEKMKVSLIETTYKLNQCFRWISSQMNRIYNVELYRYI 659
 QY 661 CDTGAFVQPALYEAFLGTVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 720
 DB 660 CDTGAFVQPALYEAFLGTVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 719
 QY 721 DLVDFPEKCKDSDHMDKISQGLKRIEKKYTWKIXSERLLTTLTGYYGFMKAVSNLERR 780
 DB 720 DLVDFPEKCKVESHMDKISQGLQRIEKKYTWKIXSERLLTTLTGYYGFMKAVSNLERR 779
 QY 781 ESRRYLEMFYALKYRKLAESVPLAE 806
 DB 780 ESRRYLEMFYALKYRKLAESVPLAE 805

RESULT 9

ID Q9XG65 PRELIMINARY; PRT; 805 AA.
 AC Q9XG65: 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Sucose synthase (EC 2.4.1.13).
 GN SUCS1.
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Chori;
 RX MEDLINE=99254811; PubMed=10323232;
 RA Hohnjec N., Becker J.D., Pehler A., Perllick A.M., Kuester H.,
 RT "Genomic organisation and expression properties of the MtsucS1 gene,
 RT which encodes a nodule-enhanced sucrose synthase from the model legume
 RT Medicago truncatula.";
 RL Mol. Gen. Genet. 261:514-522 (1999).
 DR EMBL; AJ131964; CAB40795.1; -.
 DR GO; GO:0016157; F:sucrose synthase activity; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR00368; Sucrose_synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucrose_synth; 1.

GN Glycosyltransferase; Transferase.
SQ SEQUENCE 805 AA; 92335 MW; 2126016E95DA29D CRC64;

Query Match 85.6%; Score 3637.5; DB 10; Length 805;
Best Local Similarity 85.5%; Pred. No. 4.8e-258;
Matches 689; Conservative 53; Mismatches 63; Indels 1; Gaps 1;

QY 1 MAERALTIVHSLSRELDLTLAHRNEILALLSRIGKGGKIIQHIIIEFAIPENRK 60
DB 1 MATELTIVHSLSRELDLTLAHRNEILALLSRIGKGGKIIQHIIIEFAIPENRK 60
QY 61 KLNGAFEEVLKASQDAIYLPVVALAVPRPGWEYIRVNVHVALVEELTVAEYLHFE 120
DB 61 KLNGAFEEVLKASQDAIYLPVVALAVPRPGWEYIRVNVHVALVEELTVAEYLHFE 120
QY 121 ELVDSSNGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 180
DB 121 ELVDSSNGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 180
QY 121 ELVDSSANGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 180
DB 121 ELVDSSANGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 180
QY 181 FLRVHCHGKMMMLNDRIQNLALQHVLRKAEYIGTLPPETPCAEFEHRFOELGERGW 240
DB 181 FLRVHCHGKMMMLNDRIQNLALQHVLRKAEYIGTLPPETPCAEFEHRFOELGERGW 240
QY 241 GDTAERVLEMIQLDLLEATDPTCEKFLGRIPVFNVNIILTPSGYFAQDNVLYGPTG 300
DB 241 GDTAERVLEMIQLDLLEATDPTCEKFLGRIPVFNVNIILTPSGYFAQDNVLYGPTG 300
QY 301 GGVVYILDOVRALLENBMLIRIKOQGLNITPRILITRLLPDAGVTCGGRLEKVGTEHS 360
DB 301 GGVVYILDOVRALLENBMLIRIKOQGLNITPRILITRLLPDAGVTCGGRLEKVGTEHS 360
QY 361 DILRVPRTEKGIYVKMISRFKWPVLETYEDVAHEISKEHGPDLIIIGNXSDGNV 420
DB 361 DILRVPRTEKGIYVKMISRFKWPVLETYEDVAHEISKEHGPDLIIIGNXSDGNV 420
QY 421 ASLNAKLGVTQCTTAHALEKTKYPSDSIYWKLEDEKXHFSCQPTADLFAMNHTDFIITS 480
DB 421 ASLNAKLGVTQCTTAHALEKTKYPSDSIYWKLEDEKXHFSCQPTADLFAMNHTDFIITS 480
QY 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTTEKR 540
DB 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTTEKR 540
QY 541 RLKSHPEIEDLLYTKVNEEHLCVLNDNKRPIITFMRPLDVKNLITGLVEMCGNPKLR 600
DB 541 RLKSHPEIEDLLYTKVNEEHLCVLNDNKRPIITFMRPLDVKNLITGLVEMCGNPKLR 600
QY 601 ELANLVVVGDRKESKDLSEKAEKMKFELIDKYNLNGQFRMISQNMRIINVELYRYI 660
DB 601 ELANLVVVGDRKESKDLSEKAEKMKFELIDKYNLNGQFRMISQNMRIINVELYRYI 660
QY 661 CDTKGAFVOPALYEAFLGLTVVEAMTGLPTFATCNGGPAEIIYHSGSGFNIDPYHQDAA 720
DB 661 CDTKGAFVOPALYEAFLGLTVVEAMTGLPTFATCNGGPAEIIYHSGSGFNIDPYHQDAA 720
QY 721 DILVDFEKKCKDPSHMDKISQSGKRIEKTWIKIYSERLLITLGVYGFQMGVSNLERK 780
DB 721 DILVDFEKKCKDPSHMDKISQSGKRIEKTWIKIYSERLLITLGVYGFQMGVSNLERK 780
QY 781 ESRRYLEMFYALKYRKLAEVPLAE 806
DB 781 ESRRYLEMFYALKYRKLAEVPLAE 806
QY 780 ESRRYLEMFYALKYRKLAEVPLAE 805
DB 780 ESRRYLEMFYALKYRKLAEVPLAE 805

RESULT 10

Q9LXL5 PRELIMINARY; PRT; 808 AA.
Q9LXL5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sucrose synthase-like protein.
GN F7K15_40.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euraside II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_Taxid:3702;
RN (1)
RP SEQUENCE FROM N.A.
RA Obermayer B., Ottensmeyer B., Duchemin D., Zeitler K., Mewes H.W.,
RA Rued S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.,
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL, AL33871; CAB89040.1; -
DR PIR, T49233; T49233.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose synth.
DR Pfam; PF00534; Glycos_transf_1.
DR Pfam; PF00662; Sucrose synth; 1.
SQ SEQUENCE 808 AA; 93002 MW; 39FDF33B6325CFB CRC64;

Query Match 85.1%; Score 3614.5; DB 10; Length 808;
Best Local Similarity 84.6%; Pred. No. 2.4e-256;
Matches 681; Conservative 58; Mismatches 65; Indels 1; Gaps 1;

QY 2 AERALTIVHSLSRELDLTLAHRNEILALLSRIGKGGKIIQHIIIEFAIPENRK 61
DB 4 AERVITRVHSGERLDATLVAQKNEFALLSRVAKGKIIQHIIIEFAIPENRK 63
QY 62 LANGAFEEVLKASQDAIYLPVVALAVPRPGWEYIRVNVHVALVEELTVAEYLHFE 121
DB 64 LKNGAFEEVLKASQDAIYLPVVALAVPRPGWEYIRVNVHVALVEELTVAEYLHFE 123
QY 122 ELVDSSNGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 181
DB 124 LVDGKNGNFVLELDFEPFNSSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMPLLE 183
QY 182 LRVHCHGKMMMLNDRIQNLALQHVLRKAEYIGTLPPETPCAEFEHRFOELGERGW 241
DB 184 LRVHCHGKMMMLNDRIQNLALQHVLRKAEYIGTLPPETPCAEFEHRFOELGERGW 243
QY 242 DTAERVLEMIQLDLLEATDPTCEKFLGRIPVFNVNIILTPSGYFAQDNVLYGPTG 301
DB 244 DTAERVLEMIQLDLLEATDPTCEKFLGRIPVFNVNIILTPSGYFAQDNVLYGPTG 303
QY 302 GGVVYILDOVRALLENBMLIRIKOQGLNITPRILITRLLPDAGVTCGGRLEKVGTEHS 361
DB 304 GGVVYILDOVRALLENBMLIRIKOQGLNITPRILITRLLPDAGVTCGGRLEKVGTEHS 363
QY 362 ILRVPRTEKGIYVKMISRFKWPVLETYEDVAHEISKEHGPDLIIIGNXSDGNV 421
DB 364 ILRVPRTEKGIYVKMISRFKWPVLETYEDVAHEISKEHGPDLIIIGNXSDGNV 422
QY 421 SILNAKLGVTQCTTAHALEKTKYPSDSIYWKLEDEKXHFSCQPTADLFAMNHTDFIITS 481
DB 423 SILNAKLGVTQCTTAHALEKTKYPSDSIYWKLEDEKXHFSCQPTADLFAMNHTDFIITS 482
QY 482 FOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTTEKR 541
DB 483 FOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTTEKR 542
QY 542 LKSHPEIEDLLYTKVNEEHLCVLNDNKRPIITFMRPLDVKNLITGLVEMCGNPKLR 601
DB 543 LKSHPEIEDLLYTKVNEEHLCVLNDNKRPIITFMRPLDVKNLITGLVEMCGNPKLR 602
QY 602 LANLVVVGDRKESKDLSEKAEKMKFELIDKYNLNGQFRMISQNMRIINVELYRYI 661
DB 603 LANLVVVGDRKESKDLSEKAEKMKFELIDKYNLNGQFRMISQNMRIINVELYRYI 662
QY 662 DTKGAFVOPALYEAFLGLTVVEAMTGLPTFATCNGGPAEIIYHSGSGFNIDPYHQDAA 721
DB 662 DTKGAFVOPALYEAFLGLTVVEAMTGLPTFATCNGGPAEIIYHSGSGFNIDPYHQDAA 721

Db	Qy	Db	Qy
663	DTKGAFQOPALYPAFGLITVEAMTCSLPFAACNGSPHEIIYHKSQSGHIDPHYGDQAAE	722	ILVDFEKKCKDDSHMDKISQGGKRIEKEYTWKLYSERLLITLGVGVGFWKHSNTERRE
722	SLADFEPKCGHDSHMDQISLGGIERIOEKYTWQISORLLITLGVGVGFWKHSNIDRLE	782	SRRYLEMFYALKRYKLAESVPLAAE
782	SRRYLEMFYALKRYKLAESVPLAAE	806	SRRYLEMFYALKRYPLQAQVPLAAE
807	SRRYLEMFYALKRYPLQAQVPLAAE	807	

RESULT 11	
Q9AVR8	
ID Q9AVR8	PRELIMINARY;
8	PRT; 804 AA

DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Sucrose synthase isoform 3 (EC 2.4.1.14).
GN SUS3.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustroids 1; Fabales; Fabaceae; Papilionoideae; Viciales; Pisum.
OX NCBI_TaxID=3888;
[1]
RN RN
RP
RP SEQUENCE FROM N.A.
RC
RC TISSUE=Leaf;
RA Barratt P., Barber L., Edwards A., Kruger N., Smith A.M., Wang T.,
RA Martin C.;
RT "Multiple Distinct Isoforms of Sucrose Synthase in the Pea Plant.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AJ311146; CAC32462.1; -
DR GO; GO:0046524; F:sucrose-phosphate synthase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycoyl. .; IEA
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose synth; 1.
KM Glycoyltransferase; Transferase
QO SEQUENCE 804 AA; 91941 MW; 6669B5D4E7A370DB CRC64;

Query Match	83.3%;	Score 3536.5;	DB 10;	Length 804;
Best Local Similarity	83.3%;	Pred. No. 1.2e-250;		
Matches 668;	Conservative 66;	Mismatches 67;	Indels 1;	Gaps 1

Qy	5	ALTRVSHLRRLDETTLLAHNEILLASTIEBKGGIILQHOOILIEFAIPAEENRKILAN	64
Dp	3	SLTHSTSLRQRPDETTLLAHNEILLSLRIIAKGGIILQHOOILIEFAEPIEENRKILAN	62
Qy	65	GAPEVLKASQOEMILVLRPMVALAVRRPGWEYIRVNTHALVVELTYLAETLHKKEELND	124
Dp	63	GVPEEVLRSQOEMILVLRPMVALAVRRPGWEYELRDVHGLVDELISAETLKKEELIVE	122
Qy	125	GSSNGNEVLLEDEEPNSSFPREPLTSSKISGNGVEFLNHLISAKLPHDESNMPLLEFLRV	184
Dp	123	GSSNENFVLELDEPEFPNASTPKPTQKISGNGVEFLNHLISAKLPHKESILQPLLEFLRL	182
Qy	185	HCHKGKMMILNDITQUNLALQNYLRAAEYLGTLPEPTPCAEEHRRPOELGREGMDTA	244
Dp	183	HNHGKTIWMDRIQUNLSLQYHLAKRAEYLLIKLAEPTYSSEBKFPOLGELREGMDTA	242
Qy	245	ERVLEMITQLLLDLLEATDPCLEKEKFGIRIPMVENVILTPHGFQAQDNVLGYPDTGOVV	304
Dp	243	ERVETIQLLLDLLDGDPDPTLETLFGIRIPMVENVILSPHGFQAQDNVLGYPDTGOQIV	302
Qy	305	YIIDQVAALEEMILIRIKOQGLNITRILILIRLLPDVAGTTCGGRLEKUYGTHSNIILR	364
Dp	303	YIIDQVAALEEMILIRIKOQGLNITRILILIRLLPDVAGTTCGGRLEKUYNTHCHILR	362

Qy	365	VPRRTKGIVRKIRIFPEKPMVYLETTYEDVAHEISKEIMHTPTLIIGNSDGNIVASL	424
Db	363	VPRRTKGIVRKIRISFPE-VMPYLETSEBDVANELAKELQKDPDLIVGNSDGNIVASLL	421
Qy	425	AHRGLVTOCTIAHALEKTKYDPDSDIYWKLEKDKYHFSQCFATDLFAMNHTDFTITSTFOE	484
Db	422	AHRGLVTOCTIAHALEKTKYSPESDIIYWKKFPDKKHFSSQFPADLFAMNHTDFTITSTFOE	481
Qy	485	IASKDTVOGVSASHAFTLPGLYRVVHIDIVDPDKFNIVSGAGAMELYFPYTEEKRLKH	544
Db	482	IASKDTVOGVSASHAFTLPGLYRVVHIDIVDPDKFNIVSGAALSTIYFPYTEERKLIS	541
Qy	545	FHEBIEIDLTYTKEVENEHLCVLNDNRNKLFTFMRPLDRVKNLTLVWCNGKPKYRLAN	604
Db	542	FHHDIEILLYSIVENBEHICVLKORSKIIITFMARLDRVKNIITGLVBCYGMAARLRLVN	601
Qy	605	LTVVGGDRRKESKDJBKAKEMKKKPELIDKTNLNGQFMJISSQMMRIINVELYRIYIDTK	664
Db	602	LTVVADDRRKESKDJBELIEMKKMGILLIETKYLNGQFMJISSQMDRIIRNBLRYVIDTK	661
Qy	665	GAFVQBALVEARGLTVVEAMTCGLPTFATCNGGPAEIIYVHKSQGFNIIDPYHGOADAAILV	724
Db	662	GAFVQPAIYEAHGLYIEMSCGLTPTATCNGGPAEIIYVHKSQGYHIDPYHGOADAATLV	721
Qy	725	DFEFCKCKDPSHWDKISQGLRIIEKXTWKIYSEBLLITLTVYGFWMKAVSNLERRRSR	784
Db	722	EFEFKSKADPTWMDKISHGLKRIIEKXTWOIYSDRLITLTVGVGFWMKAVTNLERRRSKR	781
Qy	785	YLEMFPALYKRLKLAESVPLAEE	806
Db	782	YLEMFPALYKSLAESVPLAVE	803

RESULT 12	
082693	
ID 082693	PRELIMINARY; PRT; 805 AA

```

DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13).
GN SUS2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Chengappa S., Loader N., Shields R.;
RT "Cloning and expression of a second tomato (Lycopersicon esculentum)
RL sucrose synthase gene.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011535; CA093681.1; -.
DR GO; GO:0016157; P:sucrose synthase activity; IEA.
DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009058; P:liposynthesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR003368; Sucrose synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose_synth; 1.
DR Glycoyltransferase; Transferase.
KW SEQUENCE 805 AA; 92465 MW; E6879774C80092E8 CRC64;

```

	Query Match	82.3%;	Score 3494.5;	DB 10;	Length 805;
	Best Local Similarity	81.8%;	Pred. No. 1.5e-247;		
	Matches	659;	Conservative	68;	Mismatches 78; Indels 1; Gaps 1
Cy	1	MAEALRVHSLARRLDETLLAARNETALLSRLEGCKGLQLQHLLEPAAPENRK	60		
Db	1	MAEEVLTRVHSLRRVDATLLAARNEILLFSTRLESCKGKGLLPKEPLAEFPALAQDDCD	60		

DT 01-OCT-2003 (TRENBLREL. 25, last annotation update)
 DE Sucrose synthase 2.
 OS Solanum tuberosum (potato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC Lamiales; Solanales; Solanaceae; Solanum.
 NCBI_taxid=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:cv. desiree;
 RA Loureiro M.E., Kopka J., Fernie A., Mueller-Roeber B., Trethewey R.;
 RT "Characterization of the phenocypases resulting specific reduction in
 RT gene expression of two sucrose synthase genes in transgenic potato
 RT plants.";
 RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY205084; AAC34668.1; -.
 DR GO: GO:0009058; P:biosynthesis; IEA.
 DR GO: GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR00368; Sucrose synth.
 DR Pfam: PF00534; Glycos_transf_1.
 DR Pfam: PF00862; Sucrose synth.
 DR SEQUENCE 803 AA; 92607 MW; 6FB50CD2E4D3162 CRC64;
 SQ
 Query Match 81.7%; Score 3470.5; DB 10; Length 803;
 Best Local Similarity 81.8%; Pred. No. 8.7e-246;
 Matches 659; Conservative 65; Mismatches 81; Indels 1; Gaps 1;
 QY 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 DB 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 QY 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 DB 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 QY 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 180
 DB 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 180
 QY 121 ELVNSTSDNPFLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 180
 DB 121 ELVNSTSDNPFLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 180
 QY 181 FLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERGW 240
 DB 181 FLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERGW 240
 QY 181 FLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERGW 240
 DB 181 FLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERGW 240
 QY 241 GPTARVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTDG 300
 DB 241 GPTARVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTDG 300
 QY 241 GPTARVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTDG 300
 DB 241 GPTARVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTDG 300
 QY 301 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 360
 DB 301 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 360
 QY 301 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 360
 DB 301 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 360
 QY 361 DILRVFPRTKEGIVRWKISREFKWPYLETTEEDVAHEISKLAGHPDILITGNXSDGIV 420
 DB 361 DILRVFPRTKEGIVRWKISREFKWPYLETTEEDVAHEISKLAGHPDILITGNXSDGIV 420
 QY 361 DILRVFPRTKEGIVRWKISREFKWPYLETTEEDVAHEISKLAGHPDILITGNXSDGIV 420
 DB 361 DILRVFPRTKEGIVRWKISREFKWPYLETTEEDVAHEISKLAGHPDILITGNXSDGIV 420
 QY 421 ASLLHKGAVTQCTTAHALEKTKYPSDLYWKKLEKHYFSCGFADLEPAMHPTFITTS 480
 DB 421 ASLLHKGAVTQCTTAHALEKTKYPSDLYWKKLEKHYFSCGFADLEPAMHPTFITTS 480
 QY 421 ASLLHKGAVTQCTTAHALEKTKYPSDLYWKKLEKHYFSCGFADLEPAMHPTFITTS 480
 DB 421 ASLLHKGAVTQCTTAHALEKTKYPSDLYWKKLEKHYFSCGFADLEPAMHPTFITTS 480
 QY 481 TFQETAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPYTEERK 540
 DB 481 TFQETAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPYTEERK 540
 QY 481 TFQETAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPYTEERK 540
 DB 481 TFQETAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPYTEERK 540
 QY 541 RLTFHPEIEDLLFSVENEHHLCVLAKDNKPILITFMARLDVKNLTGLVEYAKNPRLR 599
 DB 541 RLTFHPEIEDLLFSVENEHHLCVLAKDNKPILITFMARLDVKNLTGLVEYAKNPRLR 599
 QY 541 RLTFHPEIEDLLFSVENEHHLCVLAKDNKPILITFMARLDVKNLTGLVEYAKNPRLR 599
 DB 541 RLTFHPEIEDLLFSVENEHHLCVLAKDNKPILITFMARLDVKNLTGLVEYAKNPRLR 599
 QY 601 ELANIVVVGSDRRKESKDLLEKAKMKMPELIDKXNLNGOPRISOMNRINVELYRI 660
 DB 601 ELANIVVVGSDRRKESKDLLEKAKMKMPELIDKXNLNGOPRISOMNRINVELYRI 660
 QY 601 ELANIVVVGSDRRKESKDLLEKAKMKMPELIDKXNLNGOPRISOMNRINVELYRI 660
 DB 601 ELANIVVVGSDRRKESKDLLEKAKMKMPELIDKXNLNGOPRISOMNRINVELYRI 660
 QY 661 CDTKAFVQPALYEAFGLTIVEAMTCGLPTFATCNGSPAEIIVHGKSGFNIDPYHQDAA 720
 DB 661 CDTKAFVQPALYEAFGLTIVEAMTCGLPTFATCNGSPAEIIVHGKSGFNIDPYHQDAA 720

DB 660 ADTRKAFVQPALYEAFGLTIVEAMTCGLPTFATCNGSPAEIIVHGKSGFNIDPYHQDAA 720
 QY 721 DILVDFPEKCKDPSHMDKISOGGLKRIEKKYTWKIVSEBLLTLGYGYEFKAVNSERR 780
 DB 720 DILVDFPEKCKDPSHMDKISOGGLKRIEKKYTWKIVSEBLLTLGYGYEFKAVNSERR 780
 QY 781 ESRRYLEMFYALKRYKLAESVPLAE 806
 DB 780 EIRRYLEMFYALKRYKLAESVPLAE 806
 RESULT 15
 ID Q9LMB7 PRELIMINARY; PRT; 803 AA.
 AC Q9LMB7;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, last annotation update)
 DE Sucrose synthase (EC 2.4.1.13).
 GN CS51.
 OS Chenopodium rubrum (Red goosefoot) (Pigweed).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Caryophyllales; Amaranthaceae; Chenopodium.
 NCBI_taxid=3560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Godt D.E., Riegel A., Roitsch T.;
 RT "Regulation of sucrose synthase expression in chenopodium rubrum:
 RT characterization of sugar induced expression in photoautotrophic
 RT suspension cultures and sink specific expression in plants.";
 RL J. Plant Physiol. 146:231-238 (1995).
 DR EMBL: X82504; CA57881.1; -.
 DR GO: GO:0016157; P:transferase activity; IEA.
 DR GO: GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
 DR GO: GO:0009058; P:biosynthesis; IEA.
 DR GO: GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR00368; Sucrose synth.
 DR Pfam: PF00534; Glycos_transf_1.
 DR Pfam: PF00862; Sucrose synth.
 DR Glycoyltransferase; transferase.
 DR SEQUENCE 803 AA; 92066 MW; 467F61D92D2FB8C CRC64;
 SQ
 Query Match 81.3%; Score 3452; DB 10; Length 803;
 Best Local Similarity 81.4%; Pred. No. 2e-244;
 Matches 657; Conservative 74; Mismatches 70; Indels 6; Gaps 5;
 QY 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 DB 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 QY 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 DB 1 MAERALTIVHSIRERLDETLAHRNEILLALSRIGKGGKILQHQLLEPFAIPEENRK 60
 QY 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 DB 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 QY 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 DB 61 KLANCAFEVYLKASQEAIVLPWVALAVPRPGWEYIRVNVNHALVVEELTVAEYLHPEKE 120
 QY 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 179
 DB 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 179
 QY 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 179
 DB 121 ELVDSSNGNPFVLELDEFPFNSSPPRPTLSKISGNGVEFLNRHLSAKLFHDKESMHPLE 179
 QY 178 DFLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERG 237
 DB 178 DFLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERG 237
 QY 178 DFLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERG 237
 DB 178 DFLRVHCHGKGNMMLNDRIOMLQVLRKAEYIGTLPEPTCAEFERHPEIGLERG 237
 QY 240 WGDTERVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTD 299
 DB 240 WGDTERVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTD 299
 QY 240 WGDTERVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTD 299
 DB 240 WGDTERVLEMIQLDLLEATDPCITLKEKLGRIIPWVFNVIITPPGYFAQDNLGYPTD 299
 QY 300 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 359
 DB 300 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 359
 QY 300 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 359
 DB 300 GGVVYILDOVRALENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLKEXVGTES 359

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:51:10 ; Search time 49 Seconds

(without alignments)
4627.734 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247
Sequence: 1 MAERALTRVHSLRRLDETL.....EMFYALKRYKLAESVPLAE 806

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09 NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10 NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	US-10-003-405-2	Sequence 2, Appl1
2	3763.5	88.6	805	US-10-424-599-258404	Sequence 258404, A
3	3744.5	88.2	806	US-10-425-114-54297	Sequence 54297, A
4	3743.5	88.1	805	US-10-424-599-154986	Sequence 154986, A
5	3658.5	86.1	806	US-10-424-599-154301	Sequence 154301, A
6	3655.5	86.1	811	US-10-424-599-154176	Sequence 50176, A
7	3652.5	86.0	806	US-10-424-599-154300	Sequence 154300, A
8	3652.5	86.0	811	US-10-425-114-55134	Sequence 55134, A
9	3643.5	85.8	805	US-10-137-036-77	Sequence 77, Appl1
10	3643.5	85.8	805	US-10-393-840-44	Sequence 44, Appl1
11	3643.5	85.8	805	US-10-393-840-144	Sequence 144, Appl1
12	3299	77.7	823	US-10-425-114-53296	Sequence 53296, A
13	3299	77.7	823	US-10-425-114-56696	Sequence 56696, A
14	3299	77.7	823	US-10-425-114-57139	Sequence 57139, A
15	3299	77.7	823	US-10-425-114-72813	Sequence 72813, A

16	3295	77.6	816	US-10-080-114A-7	Sequence 7, Appl1
17	3272.5	77.1	807	US-10-425-114-57840	Sequence 57840, A
18	3272.5	77.1	811	US-10-425-114-53965	Sequence 53965, A
19	3270.5	77.0	815	US-10-289-757-170	Sequence 170, Appl1
20	3268.5	77.0	802	US-10-080-114A-5	Sequence 5, Appl1
21	3267	76.9	814	US-10-289-757-76	Sequence 76, Appl1
22	3254.5	76.6	808	US-10-289-757-169	Sequence 169, Appl1
23	3250.5	76.5	808	US-10-289-757-74	Sequence 74, Appl1
24	3248	76.5	816	US-10-289-757-77	Sequence 77, Appl1
25	3234.5	76.2	808	US-10-289-757-75	Sequence 75, Appl1
26	3158.5	74.4	670	US-10-425-114-49869	Sequence 49869, A
27	3127.5	73.6	749	US-10-425-114-56196	Sequence 56196, A
28	3005	70.8	824	US-10-425-114-45934	Sequence 45934, A
29	2892	70.4	812	US-10-424-599-241731	Sequence 241731, A
30	2872	70.0	809	US-10-080-114A-12	Sequence 12, Appl1
31	2861	69.7	802	US-10-080-114A-2	Sequence 2, Appl1
32	2914.5	68.6	619	US-10-425-114-68121	Sequence 68121, A
33	2875	67.7	797	US-10-217-939-30	Sequence 30, Appl1
34	2866.5	67.5	805	US-10-217-939-28	Sequence 28, Appl1
35	2826.5	66.6	593	US-10-425-114-50109	Sequence 50109, A
36	2676.5	63.0	652	US-10-425-114-46682	Sequence 46682, A
37	2675.5	63.0	639	US-10-425-114-58450	Sequence 58450, A
38	2552.5	60.1	596	US-10-425-114-62525	Sequence 62525, A
39	2484	58.5	651	US-10-425-114-49911	Sequence 49911, A
40	2318.5	54.6	570	US-10-425-114-54291	Sequence 54291, A
41	2286.5	53.8	527	US-10-425-114-62319	Sequence 62319, A
42	2254	53.1	567	US-10-424-599-241717	Sequence 241717, A
43	2177.5	51.3	511	US-10-425-114-58452	Sequence 58452, A
44	2138.5	50.4	628	US-10-424-599-244151	Sequence 244151, A
45	2137.5	50.3	503	US-10-425-114-72784	Sequence 72784, A

ALIGNMENTS

US-10-003-405-2

Sequence 1, Application US/10003405

Publicat ion No. US20020116736A1

GENERAL INFORMATION:

APPLICANT: Ruam, Yong-Ling

APPLICANT: Fuhbank, Robert T.

APPLICANT: Danny, Llewellyn J.

TITLE OF INVENTION: Modification of sucrose synthase gene expression in plant tissue

TITLE OF INVENTION: uses therefor

FILE REFERENCE: GHSUSY WO1

CURRENT APPLICATION NUMBER: US/10/003, 405

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: 60/251852

PRIOR FILING DATE: 2000-12-08

NUMBER OF SEQ ID NOS: 6

SOFTWARE: Patent in version 3.1

SEQ ID NO 2

LENGTH: 806

TYPE: PRT

ORGANISM: Gossypium hirsutum

FEATURE:

NAME/KEY: misc feature

LOCATION: (414)..(414)

OTHER INFORMATION: The 'Xaa' at location 414 stands for Asn, Asp, His, or Tyr.

NAME/KEY: misc feature

LOCATION: (1240)..(1240)

OTHER INFORMATION: n = any nucleotide (a,g,c,t)

US-10-003-405-2

Query Match 100.0%; Score 4245; DB 13; Length 806;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAERALTRVHSLRRLDETLAHRNEITLALSRIGKKGILQHQIILFEAIPENRK 60

DB 1 MAERALTRVHSLRRLDETLAHRNEITLALSRIGKKGILQHQIILFEAIPENRK 60

```
QY 61 KLANGAFEEVLKASQGEAIVLPFWVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPFWVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHKE 120
QY 121 ELVNGSSNGNFVLELDEFPNSSFPRTLKSGISNGVFLNRHLSAKLPHDKESMHPLE 180
DB 121 ELVNGSSNGNFVLELDEFPNSSFPRTLKSGISNGVFLNRHLSAKLPHDKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLNALOHVLRKAEYLGTLPPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLNALOHVLRKAEYLGTLPPETPCAEFEHRFOEIGLERGW 240
QY 241 GDTAEVLVEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGPDG 300
DB 241 GDTAEVLVEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGPDG 300
QY 301 GQVVYILDOVRALNEMMLRIKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTBHS 360
DB 301 GQVVYILDOVRALNEMMLRIKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTBHS 360
QY 361 DILRVFPRTEKGIYRKWISRFKWPVLETYTEDEVAHEISKELHGTPLDIGNSDGNIV 420
DB 361 DILRVFPRTEKGIYRKWISRFKWPVLETYTEDEVAHEISKELHGTPLDIGNSDGNIV 420
QY 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOFTADLFAMNHTDFIITS 480
DB 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOFTADLFAMNHTDFIITS 480
QY 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSPGADMEIYFPYTEEKR 540
DB 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSPGADMEIYFPYTEEKR 540
QY 541 RLKHFHPEIEDLLTYKVENEBHLCVLNDRNKPIILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIEDLLTYKVENEBHLCVLNDRNKPIILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 601 ELANLVVVGDRRKESKDLBEKAKMKMPELIDKYNLNGOFRMISQMNRIANVELYRYI 660
DB 601 ELANLVVVGDRRKESKDLBEKAKMKMPELIDKYNLNGOFRMISQMNRIANVELYRYI 660
QY 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGPAEIIYHKSQFNIDPHGDDAA 720
DB 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGPAEIIYHKSQFNIDPHGDDAA 720
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLGYVGFPMKAVSNLERR 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLGYVGFPMKAVSNLERR 780
QY 781 ESRRYLEMFYALKYRKLAESVPLAE 806
DB 781 ESRRYLEMFYALKYRKLAESVPLAE 806

RESULT 2
US-10-424-599-258404
; Sequence 258404, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258404
; LENGTH: 805
; TYPE: PRT
; ORGANISM: glycine max
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_75363C.1.pep
US-10-424-599-258404
Query Match 88.6%; Score 3763.5; DB 12; Length 805;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 709; Conservative 49; Mismatches 47; Indels 1; Gaps 1;

QY 1 MAERALTNRVSLRERLDETLAHRNELIALSRIBGKGIQHQIILEEBAIPEENRK 60
DB 1 MATDRLTRVSLRERLDETLAHRNELIALSRIBGKGIQHQIILEEBAIPEENRK 60
QY 61 KLANGAFEEVLKASQGEAIVLPFWVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPFWVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHKE 120
QY 121 ELVNGSSNGNFVLELDEFPNSSFPRTLKSGISNGVFLNRHLSAKLPHDKESMHPLE 180
DB 121 ELVNGSSNGNFVLELDEFPNSSFPRTLKSGISNGVFLNRHLSAKLPHDKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLNALOHVLRKAEYLGTLPPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLNALOHVLRKAEYLGTLPPETPCAEFEHRFOEIGLERGW 240
QY 241 GDTAEVLVEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGPDG 300
DB 241 GDTAEVLVEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGPDG 300
QY 301 GQVVYILDOVRALNEMMLRIKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTBHS 360
DB 301 GQVVYILDOVRALNEMMLRIKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTBHS 360
QY 361 DILRVFPRTEKGIYRKWISRFKWPVLETYTEDEVAHEISKELHGTPLDIGNSDGNIV 420
DB 361 DILRVFPRTEKGIYRKWISRFKWPVLETYTEDEVAHEISKELHGTPLDIGNSDGNIV 420
QY 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOFTADLFAMNHTDFIITS 480
DB 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOFTADLFAMNHTDFIITS 480
QY 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSPGADMEIYFPYTEEKR 540
DB 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSPGADMEIYFPYTEEKR 540
QY 541 RLKHFHPEIEDLLTYKVENEBHLCVLNDRNKPIILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
DB 541 RLKHFHPEIEDLLTYKVENEBHLCVLNDRNKPIILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 601 ELANLVVVGDRRKESKDLBEKAKMKMPELIDKYNLNGOFRMISQMNRIANVELYRYI 660
DB 601 ELANLVVVGDRRKESKDLBEKAKMKMPELIDKYNLNGOFRMISQMNRIANVELYRYI 660
QY 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGPAEIIYHKSQFNIDPHGDDAA 720
DB 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGPAEIIYHKSQFNIDPHGDDAA 720
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLGYVGFPMKAVSNLERR 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLGYVGFPMKAVSNLERR 780
QY 781 ESRRYLEMFYALKYRKLAESVPLAE 806
DB 781 ESRRYLEMFYALKYRKLAESVPLAE 806

RESULT 3
US-10-425-114-54297
; Sequence 54297, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
```

; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53333)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 54297
 ; LENGTH: 806
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700852943_Fli.pep
 ; US-10-425-114-54297

Query Match 88.2%; Score 3744.5; DB 12; Length 806;
 Best Local Similarity 87.5%; Pred. No. 0;
 Matches 705; Conservative 51; Mismatches 49; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRERLDELTLAHREILALSRIGKGILOHIOIIEPAIPEENRK 60
 DB 2 MATDRITVHSRERLDELTLAHREILALSRIGKGILOHIOIIEPAIPEENRK 61
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
 DB 62 KLTDGAFGEVLRSIOEALVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 121
 QY 121 ELVDSSNGNFVLELDEFPNSPPRPTLSKISGVGEFLNRHLSAKLFHDKESNHPLE 180
 DB 122 ELVDSSNGNFVLELDEFPNSPPRPTLSKISGVGEFLNRHLSAKLFHDKESNHPLE 181
 QY 181 FLRVCHGKMMNDRIQNALOHVLRKAEYIGTLPETPCAEFEHRPOEIGLERGW 240
 DB 182 FLRLHSVKGKTLMDRIQNSDALQHVLRKAEYIGTLPETPCAEFEHRPOEIGLERGW 241
 QY 241 GDTAERVLKEMIQLLDLLEATDPTLEKFLGRIPWVFNVILTPHGYFAODNVLGYPDTG 300
 DB 242 GDMARVLESIOQLDLLEAEDPCTLETFGRIPWVFNVILTPHGYFAODNVLGYPDTG 301
 QY 301 GQVYVILDOVALENEMLRIRKQGLDIPRILITRLLPDAIGTTCGRLEKVGTEHS 360
 DB 302 GQVYVILDOVALENEMLRIRKQGLDIPRILITRLLPDAIGTTCGRLEKVGTEHS 361
 QY 361 DILRVFTEKGIYRKWISREKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
 DB 362 HILRVFTEKGIYRKWISREKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
 QY 421 ASLTAHKGVTQCTTAHLEKTKYPSDIPYMKLEDKYHSCOPTADLPANMHTDFITTS 480
 DB 421 ASLTAHKGVTQCTTAHLEKTKYPSDIPYMKLEDKYHSCOPTADLPANMHTDFITTS 480
 QY 481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKR 540
 DB 481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKR 540
 QY 541 RLKHFHPIEDLLTYKVENESHLCVLDRNKPILFTMRPLDRVKNLTGLVEMCGKNPKLR 600
 DB 541 RLKHFHPIEDLLTYKVENESHLCVLDRNKPILFTMRPLDRVKNLTGLVEMCGKNPKLR 600
 QY 601 ELANLVVVGSDRRKESKDLSEKAEKMKFELIDKYNLNGOPRMTSSQNNRIIRNVELYXI 660
 DB 601 ELANLVVVGSDRRKESKDLSEKAEKMKFELIDKYNLNGOPRMTSSQNNRIIRNVELYXI 660
 QY 661 CDTKGAFVOPALYEAFGTLVVEAMTCGLPTFATCNGPAEIIIVHSGSGFNIDPYHGDAA 720
 DB 661 CDTKGAFVOPALYEAFGTLVVEAMTCGLPTFATCNGPAEIIIVHSGSGFNIDPYHGDAA 720
 QY 721 DILVDFEKKCKDPSHMDKISGGGLRIEKKYTWKIYSEBLLTLTGYYGFWKHSNLERR 780
 DB 721 DILVDFEKKCKDPSHMDKISGGGLRIEKKYTWKIYSEBLLTLTGYYGFWKHSNLERR 780
 QY 781 ESRRLVEMFYALKYRKLAESVPLAE 806

DB 781 ESRRLVEMFYALKYRKLAESVPLAE 806

RESULT 4
 US-10-424-599-154986
 ; Sequence 154986, Application US/10424599
 ; Publication No. US20040031072A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 154986
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_110974C.1.pep
 ; US-10-424-599-154986

Query Match 88.1%; Score 3743.5; DB 12; Length 805;
 Best Local Similarity 87.5%; Pred. No. 0;
 Matches 705; Conservative 51; Mismatches 49; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRERLDELTLAHREILALSRIGKGILOHIOIIEPAIPEENRK 60
 DB 1 MATDRITVHSRERLDELTLAHREILALSRIGKGILOHIOIIEPAIPEENRK 60
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
 DB 61 KLTDGAFGEVLRSIOEALVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
 QY 121 ELVDSSNGNFVLELDEFPNSPPRPTLSKISGVGEFLNRHLSAKLFHDKESNHPLE 180
 DB 121 ELVDSSNGNFVLELDEFPNSPPRPTLSKISGVGEFLNRHLSAKLFHDKESNHPLE 180
 QY 181 FLRVCHGKMMNDRIQNALOHVLRKAEYIGTLPETPCAEFEHRPOEIGLERGW 240
 DB 181 FLRLHSVKGKTLMDRIQNSDALQHVLRKAEYIGTLPETPCAEFEHRPOEIGLERGW 240
 QY 241 GDTAERVLKEMIQLLDLLEATDPTLEKFLGRIPWVFNVILTPHGYFAODNVLGYPDTG 300
 DB 241 GDMARVLESIOQLDLLEAEDPCTLETFGRIPWVFNVILTPHGYFAODNVLGYPDTG 300
 QY 301 GQVYVILDOVALENEMLRIRKQGLDIPRILITRLLPDAIGTTCGRLEKVGTEHS 360
 DB 301 GQVYVILDOVALENEMLRIRKQGLDIPRILITRLLPDAIGTTCGRLEKVGTEHS 360
 QY 361 DILRVFTEKGIYRKWISREKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
 DB 361 DILRVFTEKGIYRKWISREKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
 QY 421 ASLTAHKGVTQCTTAHLEKTKYPSDIPYMKLEDKYHSCOPTADLPANMHTDFITTS 480
 DB 421 ASLTAHKGVTQCTTAHLEKTKYPSDIPYMKLEDKYHSCOPTADLPANMHTDFITTS 480
 QY 481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKR 540
 DB 481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKR 540
 QY 541 RLKHFHPIEDLLTYKVENESHLCVLDRNKPILFTMRPLDRVKNLTGLVEMCGKNPKLR 600
 DB 541 RLKHFHPIEDLLTYKVENESHLCVLDRNKPILFTMRPLDRVKNLTGLVEMCGKNPKLR 600
 QY 601 ELANLVVVGSDRRKESKDLSEKAEKMKFELIDKYNLNGOPRMTSSQNNRIIRNVELYXI 660

```
Db 600 ELVNLVAVGDRRKESKDLSEKAEKKYVGLIETFKLNGCFRMISSQMNRRVANGELRYVI 659
Qy 661 CDTKGAFVQPALYEAFGITVVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDOAA 720
Db 660 CDTKGAFVQPALYEAFGITVVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDOAA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIIEKYTWKIYSERLTLTGYYGFMKHYSNLEERR 780
Db 720 DILVDFEKKCKDPSHMDKISQGLKRIIEKYTWKIYSERLTLTGYYGFMKHYSNLEERR 779
Qy 781 ESRRYLEMFFYALKYRKLAESVPLAE 806
Db 780 ESRRYLEMFFYALKYRKLAESVPLAE 805

RESULT 5
US-10-424-599-154301
; Sequence 154301, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154301
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110355C.1.pep
US-10-424-599-154301

Query Match      86.1%; Score 3656.5; DB 12; Length 806;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 687; Conservative 59; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERALTIVHSIRERLDETLAHNEIILALSRIGKGGKILQHQIILEFPALPEERK 60
Db 1 MANHPLTSHSFRERFDETLGHRNEIILALSRLEAKGGKILQHQVAAEFEEIPEESRK 60
Qy 61 KLANGAFFEVYKASQEAIVLPWVALAVPRPGWVEYLRVNVHAIYVEELVAEYLHPKE 120
Db 61 KLOGGVFGEVLRSTQEAIVLPFVALAVPRPGWVEYLRVNVHAIYVEELVAEYLHPKE 120
Qy 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Qy 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Qy 161 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEEYIGTLPPETPCAEFHRFOEIGLENGW 240
Db 161 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEEYIGTLPPETPCAEFHRFOEIGLENGW 240
Qy 181 FLRLHSYKGTMMNDKQVSLDSLOHVLRKAEEYITSVAPETPYSEFENKEREIGLENGW 240
Db 181 FLRLHSYKGTMMNDKQVSLDSLOHVLRKAEEYITSVAPETPYSEFENKEREIGLENGW 240
Qy 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
Db 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
Qy 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
Db 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
Qy 301 GGVVYLLDQVRLAENEMILRIKQOGNITPRILITRLLPDAVGTTCQORLEKYVGTSHS 360
Db 301 GGVVYLLDQVRLAENEMILRIKQOGNITPRILITRLLPDAVGTTCQORLEKYVGTSHS 360
Qy 361 DILRVFPRTKGIYVRKMSRFEKWPYLETYEDVAHEISKELHGTPLLIGNSDGNIV 420
Db 361 DILRVFPRTKGIYVRKMSRFEKWPYLETYEDVAHEISKELHGTPLLIGNSDGNIV 420
Qy 421 ASLLAHKLGVTQCTIAHALEKTKYPDSOIYMKKLDKHYHSCGFADLFAMNHTPTITS 480
Db 421 ASLLAHKLGVTQCTIAHALEKTKYPDSOIYMKKLDKHYHSCGFADLFAMNHTPTITS 480
Qy 480 ASLLAHKLGVTQCTIAHALEKTKYPDSOIYMKKLDKHYHSCGFADLFAMNHTPTITS 480
Db 480 ASLLAHKLGVTQCTIAHALEKTKYPDSOIYMKKLDKHYHSCGFADLFAMNHTPTITS 480
```

```
Qy 481 TPQEIAGSKDTVGQYSEHTAFTLPGLYRVYHGDIVPEPKNNIYSPGADMETYPPEYTEKR 540
Db 480 TPQEIAGSKDTVGQYSEHTAFTLPGLYRVYHGDIVPEPKNNIYSPGADMETYPPEYTER 539
Qy 541 RLKHFHEIDILYTKYNESEHLCVLNDRNKPIFTMPRIIDRVNGLGLVEMCKNPKLR 600
Db 540 RLKHFHEIDILYTKYNESEHLCVLNDRNKPIFTMPRIIDRVNGLGLVEMCKNPKLR 599
Qy 601 ELANLVVVGDRRKESKDLSEKAEKKYVGLIETFKLNGCFRMISSQMNRIYANGELRYVI 660
Db 600 ELVNLVAVGDRRKESKDLSEKAEKKYVGLIETFKLNGCFRMISSQMNRRVANGELRYVI 659
Qy 661 CDTKGAFVQPALYEAFGITVVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDOAA 720
Db 660 CDTKGAFVQPALYEAFGITVVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDOAA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIIEKYTWKIYSERLTLTGYYGFMKHYSNLEERR 780
Db 720 DILVDFEKKCKDPSHMDKISQGLKRIIEKYTWKIYSERLTLTGYYGFMKHYSNLEERR 779
Qy 781 ESRRYLEMFFYALKYRKLAESVPLAE 806
Db 780 ESRRYLEMFFYALKYRKLAESVPLAE 805

RESULT 6
US-10-425-114-50176
; Sequence 50176, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50176
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700852649_F11.pep
US-10-425-114-50176

Query Match      86.1%; Score 3655.5; DB 12; Length 811;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 687; Conservative 59; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERALTIVHSIRERLDETLAHNEIILALSRIGKGGKILQHQIILEFPALPEERK 60
Db 1 MANHPLTSHSFRERFDETLGHRNEIILALSRLEAKGGKILQHQVAAEFEEIPEESRK 65
Qy 61 KLANGAFFEVYKASQEAIVLPWVALAVPRPGWVEYLRVNVHAIYVEELVAEYLHPKE 120
Db 61 KLOGGVFGEVLRSTQEAIVLPFVALAVPRPGWVEYLRVNVHAIYVEELVAEYLHPKE 125
Qy 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 185
Qy 181 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEEYIGTLPPETPCAEFHRFOEIGLENGW 240
Db 181 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEEYIGTLPPETPCAEFHRFOEIGLENGW 245
Qy 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
Db 241 GGTARVLEMIQLLDLEATDPCLEKFLGRIPWVFNVVILTPHGYFAQDNVLYGYPDTG 300
```


Db 246 GDIAERVLEMIQLLDLLEADPCTLEFLGRVNFVNLSPHGYAQNVLGYPDTG 305
Qy 301 GGVVYILDOVRLAENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYGTSHS 360
Db 306 GGVVYILDOVRLAENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYGTSHS 365
Qy 361 DILRVPFTEKGIYAKWISREKVPYLETYEDVAHEISKEHAGTDPDLIGNXSDGNIV 420
Db 366 DILRVPFTEKGIYAKWISREKVPYLETYEDVAHEISKEHAGTDPDLIGNXSDGNIV 424
Qy 421 ASLHAKIGVTOCTTAHLEKTYPDSDIYMKKLEDKTHFSQCFADLPFAMNHTDPIITS 480
Db 425 ASLHAKIGVTOCTTAHLEKTYPDSDIYMKKLEDKTHFSQCFADLPFAMNHTDPIITS 484
Qy 481 TFOETAGSKDVTGYESHTAFTLPGLYRVHGIIDPDPKFNIVSPGADMEIYFPTTEER 540
Db 485 TFOETAGSKDVTGYESHTAFTLPGLYRVHGIIDPDPKFNIVSPGADMEIYFPTTEER 544
Qy 541 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMCGNPKLR 600
Db 545 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMCGNPKLR 604
Qy 601 ELANLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 660
Db 605 ELANLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 664
Qy 661 CDTKAFQOPALYEAFTLVVAMTCGLPTFATCNGSPAEIIVHKSQFNIDPYGDQAA 720
Db 665 CDTKAFQOPALYEAFTLVVAMTCGLPTFATCNGSPAEIIVHKSQFNIDPYGDQAA 724
Qy 721 DILVDFEKKCKDPESHMDKISQGLKRIEKKYTWKISERLLTLGVYGFWMGVNLEBR 780
Db 725 DILVDFEKKCKDPESHMDKISQGLKRIEKKYTWKISERLLTLGVYGFWMGVNLEBR 784
Qy 781 ESRRLYEMFYALKYRKLAESVPLAE 806
Db 785 ESRRLYEMFYALKYRKLAESVPLAE 810

RESULT 7
US-10-424-599-154300
; Sequence 154300, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154300
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110354C.1.pep
US-10-424-599-154300

Query Match 86.0%; Score 3652.5; DB 12; Length 806;
Best Local Similarity 85.4%; Fred. No. 0;
Matches 688; Conservative 58; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERLTVRHSLRERLDELTLAHRNEIALLSRIEGKGIGIHOIILEFAIPENRX 60
Db 1 MANHLTVSHSRERFDELTLGHRNEIALLSRLEKKGIGIHOIIVAEFEIIBESRX 60
Qy 61 KLANGAFEEVLKASQEAIVLPWVALAVRPRGVWEYIRVNVALVVEELTVAAVYLFKE 120
Db 61 KLQDGVGVGLASTQDAIVLPFVALAVRPRGVWEYIRVNVALVVEELTVAAVYLFKE 120

Qy 121 ELVDSSNGNFEVLEIDFEFPNSSFPRLTSKISGNGVEFLNRHLSAKLPHDKESHAPLE 180
Db 121 ELVDSSNGNFEVLEIDFEFPNSSFPRLTSKISGNGVEFLNRHLSAKLPHDKESHAPLE 180
Qy 181 FLRVCHGKMMMLNDRIQNLALQHVLRKAEVYGLTPEPTCAEFERFOEIGLRGW 240
Db 181 FLRVCHGKMMMLNDRIQNLALQHVLRKAEVYGLTPEPTCAEFERFOEIGLRGW 240
Qy 241 GDTARVLEMIQLLDLLEADPCTLEFLGRVNFVNLSPHGYAQNVLGYPDTG 300
Db 241 GDTARVLEMIQLLDLLEADPCTLEFLGRVNFVNLSPHGYAQNVLGYPDTG 300
Qy 301 GGVVYILDOVRLAENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYGTSHS 360
Db 301 GGVVYILDOVRLAENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYGTSHS 360
Qy 361 DILRVPFTEKGIYAKWISREKVPYLETYEDVAHEISKEHAGTDPDLIGNXSDGNIV 420
Db 361 DILRVPFTEKGIYAKWISREKVPYLETYEDVAHEISKEHAGTDPDLIGNXSDGNIV 420
Qy 421 ASLHAKIGVTOCTTAHLEKTYPDSDIYMKKLEDKTHFSQCFADLPFAMNHTDPIITS 480
Db 421 ASLHAKIGVTOCTTAHLEKTYPDSDIYMKKLEDKTHFSQCFADLPFAMNHTDPIITS 480
Qy 481 TFOETAGSKDVTGYESHTAFTLPGLYRVHGIIDPDPKFNIVSPGADMEIYFPTTEER 540
Db 481 TFOETAGSKDVTGYESHTAFTLPGLYRVHGIIDPDPKFNIVSPGADMEIYFPTTEER 540
Qy 541 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMCGNPKLR 600
Db 541 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMCGNPKLR 600
Qy 601 ELANLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 660
Db 601 ELANLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 660
Qy 660 ELVNLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 664
Db 660 ELVNLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 664
Qy 661 CDTKAFQOPALYEAFTLVVAMTCGLPTFATCNGSPAEIIVHKSQFNIDPYGDQAA 720
Db 661 CDTKAFQOPALYEAFTLVVAMTCGLPTFATCNGSPAEIIVHKSQFNIDPYGDQAA 720
Qy 721 DILVDFEKKCKDPESHMDKISQGLKRIEKKYTWKISERLLTLGVYGFWMGVNLEBR 780
Db 721 DILVDFEKKCKDPESHMDKISQGLKRIEKKYTWKISERLLTLGVYGFWMGVNLEBR 780
Qy 780 ELVNLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 664
Db 780 ELVNLVVVGDRRKSSKDLBEKAEKMKFELIDKTNLNGQFRWISSQMRINVELYRYI 664

RESULT 8
US-10-425-114-55184
; Sequence 55184, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313) B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55184
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700565776_F11.pep

US-10-425-114-55184

Query Match 86.0%; Score 3652.5; DB 12; Length 811;
 Best Local Similarity 85.4%; Pred. No. 0;
 Matches 688; Conservative 58; Mismatches 59; Indels 1; Gaps 1;

```

QY 1 MAERALTAVHSIRERLDETLAHNRNEIALLSRIEGKGGILOHHQIILEFEALPEENRK 60
DB 6 MANHPLTSHSFRERFDETLGHRNEIALLSRLBAKGKGILOHHQVAAFEFEIPEESRK 65
QY 61 KLANGAFFEVUKAQSQAIVLPBWVALAVPRPGWMEYIRUNVHALVVELVAEYLHKE 120
DB 66 KLQDVFGEVLRKSTQEAIVLPFWALAVPRPGWMEYIRUNVHALVVELVAEYLHKE 125
QY 121 ELVDSSNGNFVLELDFEPFNSFPRLTSGISNGVEFLNRHLSAKLFHDKESHAPLE 180
DB 126 ELVESSNGNFVLELDFEPFNSFPRLTSGISNGVEFLNRHLSAKLFHDKESHAPLE 185
QY 181 FLRVHCHGKMMNDRIQINMALQHVLRKAEYLGTLRPETPCAEFEHRFOEIGLERGW 240
DB 186 FLRLHSYKGTWMLNDKQSLDSLOHVLARKAEYLSVAPETPYSEFENRFRIGLERGW 245
QY 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPWVFNVIITPBGYFAQDNVLAGYPTG 300
DB 246 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPWVFNVIITPBGYFAQDNVLAGYPTG 305
QY 301 GOVVYILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDVAGTTCGRLEKVGTEHS 360
DB 306 GOVVYILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDVAGTTCGRLEKVGTEHS 365
QY 361 DILRVPFTEKGIYRKWISREKWPYLETYTEVAHEISKELHGTPLDIIGNSDGNIV 420
DB 366 DILRVPFTEKGIYRKWISREKWPYLETYTEVAHEISKELHGTPLDIIGNSDGNIV 424
QY 421 ASLNAHKIYVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFADLPANMHTDFITTS 480
DB 425 ASLNAHKIYVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFADLPANMHTDFITTS 484
QY 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 540
DB 485 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 544
QY 541 RLKHFHPEIEDLLYTKVENEHLCVYLANDRNPILFTMPRLDRVXNLJGLVEMCGNPKLR 600
DB 545 RLKHFHPEIEDLLYTKVENEHLCVYLANDRNPILFTMPRLDRVXNLJGLVEMCGNPKLR 604
QY 601 ELANLVVVGGRRRKSKDLBEKAEMKKMFELIDKYNLNGQFRWISSQNNRIRNVLYRYI 660
DB 605 ELVNLVVVAAGDRRKSKDLBEKAEMKKMFELIDKYNLNGQFRWISSQNNRIRNVLYRYI 664
QY 661 CDTKGAFOPALYEAFGTLVVEAMTQGLPTPATCGGPAEIVHGKSGFNIDPYHGDQA 720
DB 665 CDTKGAFOPALYEAFGTLVVEAMTQGLPTPATCGGPAEIVHGKSGFNIDPYHGDQA 724
QY 721 DILVDFEKKCKDPESHMDKISQGLKRIIEKXYTWKISERLLTLTGYYGFWMGRVSNLERR 780
DB 725 ELVVEFEKSKADPSHMDKISQGLKRIIEKXYTWKISERLLTLTGYYGFWMGRVSNLERR 784
QY 781 ESRRLVEMFYALKYRKLAEVPLAE 806
DB 785 ESKRYLEMFYALKYRKLAEVPLAE 810

```

RESULT 9

US-10-137-036-77

; Sequence 77, Application US/10137036

; Publication No. US20030101478A1

; GENERAL INFORMATION:

; APPLICANT: Perera, Ranjan

; APPLICANT: Rice, Stephen

; APPLICANT: Bagleton, Claire

; APPLICANT: Lasham, Annette

; APPLICANT: Wood, Marion

```

; APPLICANT: Visser, Elizabeth
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1036c4
; CURRENT APPLICATION NUMBER: US/10/137,036
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: PCT/NZ00/00018
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Eucalyptus grandis

```

US-10-137-036-77

Query Match 85.8%; Score 3643.5; DB 14; Length 805;
 Best Local Similarity 85.0%; Pred. No. 0;
 Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

```

QY 1 MAERALTAVHSIRERLDETLAHNRNEIALLSRIEGKGGILOHHQIILEFEALPEENRK 60
DB 1 MADRWLTSHSLRERLDETLAHNRNDIYAFISVBAKGKGILOHHQIIPAEFEALSEESRA 60
QY 61 KLANGAFFEVUKAQSQAIVLPBWVALAVPRPGWMEYIRUNVHALVVELVAEYLHKE 120
DB 61 KLQDGAFFEVUKSTQEAIVSPWVALAVPRPGWMEYIRUNVHALVVELVAEYLHKE 120
QY 121 ELVDSSNGNFVLELDFEPFNSFPRLTSGISNGVEFLNRHLSAKLFHDKESHAPLE 180
DB 121 ELADGSLNGNFVLELDFEPFNSFPRLTSGISNGVEFLNRHLSAKLFHDKESHAPLE 180
QY 181 FLRVHCHGKMMNDRIQINMALQHVLRKAEYLGTLRPETPCAEFEHRFOEIGLERGW 240
DB 181 FLQVHCHGKMMNVNARIQNVFSLQHVLRKAEYLTSLKPEYPSQFHNKFOEIGLERGW 240
QY 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPWVFNVIITPBGYFAQDNVLAGYPTG 300
DB 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPWVFNVIITPBGYFAQDNVLAGYPTG 300
QY 301 GOVVYILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDVAGTTCGRLEKVGTEHS 360
DB 301 GOVVYILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDVAGTTCGRLEKVGTEHS 360
QY 361 DILRVPFTEKGIYRKWISREKWPYLETYTEVAHEISKELHGTPLDIIGNSDGNIV 420
DB 361 HILRVPFNEKGVYRKWISREKWPYLETYTEVAHEISKELHGTPLDIIGNSDGNIV 419
QY 421 ASLNAHKIYVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFADLPANMHTDFITTS 480
DB 420 ASLNAHKIYVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFADLPANMHTDFITTS 479
QY 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 540
DB 480 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 539
QY 541 RLKHFHPEIEDLLYTKVENEHLCVYLANDRNPILFTMPRLDRVXNLJGLVEMCGNPKLR 600
DB 540 RLKHFHPEIEDLLYTKVENEHLCVYLANDRNPILFTMPRLDRVXNLJGLVEMCGNPKLR 599
QY 601 ELANLVVVGGRRRKSKDLBEKAEMKKMFELIDKYNLNGQFRWISSQNNRIRNVLYRYI 660
DB 600 ELANLVVVGGRRRKSKDLBEKAEMKKMFELIDKYNLNGQFRWISSQNNRIRNVLYRYI 659

```

```

Qy 661 CDTKGAFOVOPALYEA FGLTVVEAMTCGLPTFATCNGGPAEIIIVHGKSGFNIDPYHGDQA 720
Db 660 CDTKGVFOVOPALYEA FGLTVVEAMTCGLPTFATCNGGPAEIIIVHGKSGHIDPYHGDQA 719
Qy 721 DILVDFPEKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTVGYGFMKHYNSLERR 780
Db 720 ELVDFPFNKCKIDQSHMDEISKAMQRIEKKYTWKYSERLLTLTVGYGFMKHYNTLDR 779
Qy 781 ESRRYLMEFVALKYRKLAEVPLAE 806
Db 780 ESRRYLMEFVALKYRKLAEVPLAE 805

RESULT 10
US-10-393-840-44
; Sequence 44, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-393-840-44

Query Match 85.8%; Score 3643.5; DB 15; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MABRALTVSHLRRLDETLAHNRNELLASRIEGKGLQHQQHILFEPAIDENRK 60
Db 1 MADMLTRSHSLRRLDETLAHNRNDIVAFLSRVEAKGKGLQHQPFAFEPAISESRA 60
Qy 61 KLANGAFPEVVKASQEAIVLPWVALAVRPPGVWEYIRVNVHVLVVEELVAEYLHFKE 120
Db 61 KLIDGAGFEVVKASQEAIVSPWVALAVRPPGVWEHIRVNVHVLVLEQLVAEYLHFKE 120
Qy 121 ELVDGSSNGNFVLELDEFPNSSFPRPTLSKISGVGFNLNRHSASLFDHKSHPLE 180
Db 121 ELADGSLNGNFVLELDEFPPTASFPPTLSKISGVGFNLNRHSASLFDHKSHPLE 180
Qy 181 FLRVHCHGKKNMMLNDRIQNLNALQHVLRKAEVYGLTPETPCAEFHRFOEIGLERGW 240
Db 181 FLQVHCHGKKNMMLNARIONVFSQHVLRKAEVYGLTPETPCAEFHRFOEIGLERGW 240
Qy 241 GDTAEVLEMTQLLDLLEATDPTLKEFGRIPVNFVNLTPHGYAONVLYGPTG 300
Db 241 GDTAEVLEMTQLLDLLEAPDPTLKEFGRIPVNFVNLTPHGYAONVLYGPTG 300
Qy 301 GQVYVILDOVALLENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKYVTEHS 360
Db 301 GQVYVILDOVALLEEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKYVTEHS 360
Qy 361 DILVFPPTTEKGIYRKMSRPEKWPVYETTEDVAHISLHGTPLDILGNXSDGNTV 420
Db 361 DILVFPPTTEKGIYRKMSRPEKWPVYETTEDVAHISLHGTPLDILGNXSDGNTV 420
Qy 421 HILFVFPNTEKGVVWKMSRPE-VWPLYERYTEDEVASELAGELQKPDLLIGNSDGNIV 419

```

```

Qy 421 ASILAHKIGVQCTIAHALEKTKYPSDSIYWKCKEDKXYHFSQCFADLIFAMHDTPIITS 480
Db 420 ASILAHKIGVQCTIAHALEKTKYPSDSIYWKCKEDKXYHFSQCFADLIFAMHDTPIITS 479
Qy 481 TFOEIASGKDVQGYESHHTAFTLPGLYRVHGDIVDFDKFNIVSDGADMEIYFYTEER 540
Db 480 TFOEIASGKDVQGYESHHTAFTLPGLYRVHGDIVDFDKFNIVSDGADMEIYFYTEER 539
Qy 541 RLKHFPEIEDLLTYKVNEBHLCYLNDNRKPIELFTMRDLRVKVLTLGVKCGKPKLR 600
Db 540 RLKHFPEIEDLLTYKVNEBHLCYLNDNRKPIELFTMRDLRVKVLTLGVKCGKPKLR 599
Qy 601 ELANLVVGGDRRDKSLEEKAEKMKFELIDKYNLNGOFWISSQNRIRNVELYRYI 660
Db 600 ELANLVVGGDRRDKSLEEKAEKMKFELIDKYNLNGOFWISSQNRIRNVELYRYI 659
Qy 661 CDTKGAFOVOPALYEA FGLTVVEAMTCGLPTFATCNGGPAEIIIVHGKSGFNIDPYHGDQA 720
Db 660 CDTKGVFOVOPALYEA FGLTVVEAMTCGLPTFATCNGGPAEIIIVHGKSGHIDPYHGDQA 719
Qy 721 DILVDFPEKCKDPSHMDKISOGGLKRIEKKYTWKYSERLLTLTVGYGFMKHYNSLERR 780
Db 720 ELVDFPFNKCKIDQSHMDEISKAMQRIEKKYTWKYSERLLTLTVGYGFMKHYNTLDR 779
Qy 781 ESRRYLMEFVALKYRKLAEVPLAE 806
Db 780 ESRRYLMEFVALKYRKLAEVPLAE 805

RESULT 11
US-10-393-840-144
; Sequence 144, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-393-840-144

Query Match 85.8%; Score 3643.5; DB 15; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MABRALTVSHLRRLDETLAHNRNELLASRIEGKGLQHQQHILFEPAIDENRK 60
Db 1 MADMLTRSHSLRRLDETLAHNRNDIVAFLSRVEAKGKGLQHQPFAFEPAISESRA 60
Qy 61 KLANGAFPEVVKASQEAIVLPWVALAVRPPGVWEYIRVNVHVLVVEELVAEYLHFKE 120
Db 61 KLIDGAGFEVVKASQEAIVSPWVALAVRPPGVWEHIRVNVHVLVLEQLVAEYLHFKE 120
Qy 121 ELVDGSSNGNFVLELDEFPNSSFPRPTLSKISGVGFNLNRHSASLFDHKSHPLE 180
Db 121 ELADGSLNGNFVLELDEFPPTASFPPTLSKISGVGFNLNRHSASLFDHKSHPLE 180
Qy 181 FLRVHCHGKKNMMLNDRIQNLNALQHVLRKAEVYGLTPETPCAEFHRFOEIGLERGW 240
Db 181 FLRVHCHGKKNMMLNDRIQNLNALQHVLRKAEVYGLTPETPCAEFHRFOEIGLERGW 240

```

Db 181 FLQVHCYKGMKMNVARIQNVFSLOHVLKRAKEEYLTSLKPETPYSQEFHKEFOEIGLBKGM 240
Qy 241 GDTREPVLEMOTLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPDG 300
Db 241 GDTREPVLEMOTLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPDG 300
Qy 301 GQVYVILDOVALENEMLLRIKQOGLNTPRILITRLLPDAVGTTCGORLEKVGTEHS 360
Db 301 GQVYVILDOVALENEMLLRIKQOGLNTPRILITRLLPDAVGTTCGORLEKVGTEHS 360
Qy 361 DILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 420
Db 361 HILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 419
Qy 421 ASLAAHKLVGTQCTTAHLEKTKYPPDSDIYWKLEDKYHFSQCFADLFAMNHTDFITTS 480
Db 420 ASLAAHKLVGTQCTTAHLEKTKYPPDSDIYWKLEDKYHFSQCFADLFAMNHTDFITTS 479
Qy 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSGADMEIYFPYTEBR 540
Db 480 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSGADMEIYFPYTEBR 539
Qy 541 RLKHPHPIEDLLYTKVNEBHLVYANDRNKPIFTMPRLDRVKULTGLVEMCGKPKLR 600
Db 540 RLKHPHPIEDLLYTKVNEBHLVYANDRNKPIFTMPRLDRVKULTGLVEMCGKPKLR 599
Qy 601 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKNLNGOFPMISSQMRIRNVELYRYIC 660
Db 600 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKNLNGOFPMISSQMRIRNVELYRYIC 659
Qy 661 CDTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGPRAEIIYHGKSGFNIDPYHGDQAA 720
Db 660 CDTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGPRAEIIYHGKSGFNIDPYHGDQAA 719
Qy 721 DILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 780
Db 720 DILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 779
Qy 781 ESRRYLEMFYALKYRKLAESVPLAEE 806
Db 780 ESRRYLEMFYALKYRKLAESVPLAEE 805

RESULT 12
US-10-425-114-53296
; Sequence 53296, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53296
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700451030_Flt.pep
US-10-425-114-53296

Query Match 77.7%; Score 3299; DB 12; Length 823;
Best Local Similarity 76.9%; Pred. No. 7, 3e-303;
Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

Qy 3 BRALTIVHSIRERDELTLAHNEILLALSRIECKGKGIQHOHIIIEF-EAIPENRKK 61
Db 14 DRLVSLRHSYREKIRIGSLSAHPNELVAVFTRLLKGLQPHQIIAEYNNALIEABREK 73
Qy 62 LANGAFPEVLKASGEAVLPPWALAVPRPGWVEYRVNVHVALVEELGTAEYLHPKEE 121
Db 74 LKQAFEDVLRQAQEAIVIPWVALAIRPRPGWVEYRVNVHVALVEELGTAEYLHPKEE 133
Qy 122 LVDSNGNENVLBELDFEPFNSPPRPPLTSKISGVGEFLNRHLSAKLPHDKESMPLLEF 181
Db 134 LVEEGPNNNFVLBELDFEPFNSPPRPPLTSKISGVGEFLNRHLSAKLPHDKESMPLLEF 193
Qy 182 LRVNCHGKMMNDRIQNALDOLVLRKAKEEYLTSPREPCCAEFHRPOELBERGKG 241
Db 194 LRAHNYGMMTMNDRIRSLALOGARKEEHLSTQADTPYSEFHRHRLQELBERGKG 253
Qy 242 DTAERVLEMIQLDLLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPDG 301
Db 254 DTAERVLEMIQLDLLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPDG 313
Qy 302 QVYVILDOVALENEMLLRIKQOGLNTPRILITRLLPDAVGTTCGORLEKVGTEHS 361
Db 314 QVYVILDOVALENEMLLRIKQOGLNTPRILITRLLPDAVGTTCGORLEKVGTEHS 373
Qy 362 ILRVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 421
Db 374 ILRVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 432
Qy 422 SILAHKLVGTQCTTAHLEKTKYPPDSDIYWKLEDKYHFSQCFADLFAMNHTDFITTS 481
Db 433 SILAHKLVGTQCTTAHLEKTKYPPDSDIYWKLEDKYHFSQCFADLFAMNHTDFITTS 492
Qy 482 FOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSGADMEIYFPYTEBR 541
Db 493 FOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSGADMEIYFPYTEBR 552
Qy 542 LKHPHPIEDLLYTKVNEBHLVYANDRNKPIFTMPRLDRVKULTGLVEMCGKPKLR 601
Db 553 LKHPHPIEDLLYTKVNEBHLVYANDRNKPIFTMPRLDRVKULTGLVEMCGKPKLR 612
Qy 602 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKNLNGOFPMISSQMRIRNVELYRYIC 661
Db 613 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKNLNGOFPMISSQMRIRNVELYRYIC 672
Qy 662 DTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGPRAEIIYHGKSGFNIDPYHGDQAA 721
Db 673 DTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGPRAEIIYHGKSGFNIDPYHGDQAA 732
Qy 722 DILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 781
Db 733 DILVFPFTEKGIYRKATISREKWPYLETETEDVAHEISKELGTPLLIGNXSDGNIV 792
Qy 782 ESRRYLEMFYALKYRKLAESVPLAEE 806
Db 793 ESRRYLEMFYALKYRKLAESVPLAEE 817

RESULT 13
US-10-425-114-56696
; Sequence 56696, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56696
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3069-052-F8_FLI.pep
US-10-425-114-56696

Query Match 77.7% Score 3299; DB 12; Length 823;
Best Local Similarity 76.9%; Pred. No. 7.3e-303;
Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

3 ERALTRVHSLEBRDELTLAHRNEILALSRLEGKGILOHQTILF-EAIPENRK 61
14 DRVLSRLHSVRERIGDSLSAHPNELVAFTRLKNGKMLQPHQIAYNNAIPEAREK 73
62 LANGAFEEVLKASQEAIVLPWVALAIPRGVWEYIRVNHALVVEELTVAEYHFKBE 121
74 LKDGAFEDVLRQAQDAIVIPWVALAIPRGVWEYIRVNHALVVEELTVAEYHFKBE 133
122 LVDSNGNPFVLELDFEPPNSSFPRPTLSKISGVNVEPLNRLSAKLPHDKSMPLLEF 181
134 LVEEGPNNPFVLELDFEPPNSSFPRPTLSKISGVNVEPLNRLSAKLPHDKSMPLLEF 193
182 LRVCHKGNMMLNDRIQNLALQHVLRKAEYIGTLPEPCAFEFERFOELGKMG 241
194 LRAHYKMTMMLNDRISLSALQALRKAEHSLTLOADTPYSEFHHRFOLGKMG 253
242 DTAERLEMIQILLDLLEATDPTLEKFLGRIIPWENVVILTPHGYFAQDNVLAGYPTGG 301
254 DCAKRAQETIHLDLLEAPDPTLEKFLGRIIPWENVVILTPHGYFAQDNVLAGYPTGG 313
302 QVVYILDOVRALLENMLRIKQGGINTPRILITRLIPDAVGTTCGRLKGVTEHSD 361
314 QVVYILDOVRALLENMLRIKQGGINTPRILITRLIPDAVGTTCGRLKGVTEHSD 373
362 ILRVFRTKGIIVRKWISRFKWPYLETYEDVAHEISKELHGTDLIGNKSDGNIVA 421
374 ILRVFRTKGIIVRKWISRFKWPYLETYEDVAHEISKELHGTDLIGNKSDGNIVA 432
422 SLAAHKGVTCTTAHALEKTKYPSDSIYMKLEDKYHSCQPTADLFAMNHTDPIITST 481
433 CLAAHKGVTCTTAHALEKTKYPSDSIYMKLEDKYHSCQPTADLFAMNHTDPIITST 492
482 FOEIAKSDTGVQYESHAFITPGLYRVVHGIDVDFPKFNIVSPGADLSIYFPYTESHR 541
493 FOEIAKSDTGVQYESHAFITPGLYRVVHGIDVDFPKFNIVSPGADLSIYFPYTESHR 552
542 LKHPPEIEDLLYTKVNEEHLCVLNDRNKPIITFMPRLDRVKLTGLVEMCGKNPKLRE 601
553 LKHPPEIEDLLYTKVNEEHLCVLNDRNKPIITFMPRLDRVKLTGLVEMCGKNPKLRE 612
602 LANLVVVGDRKESKDLLEKAEKMKFELIDKYNLNGOFWISSOMNRIRNVELYRITC 661
613 LANLVVVGDRKESKDLLEKAEKMKFELIDKYNLNGOFWISSOMNRIRNVELYRITC 672
662 DTGAFFVQPALYEAFTLVVAMTCGLPTFATCNGSPAETIVHKGSGFNIDPYHGDQAD 721
673 DTGAFFVQPALYEAFTLVVAMTCGLPTFATCNGSPAETIVHKGSGFNIDPYHGDQAD 732
722 ILVDFEKKCKDPSHMDKISQGLKRIEKTWKIYSERLLTLTGVYGFVKVSLERE 781
733 ILVDFEKKCKDPSHMDKISQGLKRIEKTWKIYSERLLTLTGVYGFVKVSLERE 792
782 SRRYLEMFAALYKRLAESVPLAE 806
793 TRRYLEMLYALKYRTMASTVPLAE 817

RESULT 14
US-10-425-114-57139
; Sequence 57139, Application US/10425114

Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57139
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700221147_FLI.pep
US-10-425-114-57139

Query Match 77.7% Score 3299; DB 12; Length 823;
Best Local Similarity 76.9%; Pred. No. 7.3e-303;
Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

3 ERALTRVHSLEBRDELTLAHRNEILALSRLEGKGILOHQTILF-EAIPENRK 61
14 DRVLSRLHSVRERIGDSLSAHPNELVAFTRLKNGKMLQPHQIAYNNAIPEAREK 73
62 LANGAFEEVLKASQEAIVLPWVALAIPRGVWEYIRVNHALVVEELTVAEYHFKBE 121
74 LKDGAFEDVLRQAQDAIVIPWVALAIPRGVWEYIRVNHALVVEELTVAEYHFKBE 133
122 LVDSNGNPFVLELDFEPPNSSFPRPTLSKISGVNVEPLNRLSAKLPHDKSMPLLEF 181
134 LVEEGPNNPFVLELDFEPPNSSFPRPTLSKISGVNVEPLNRLSAKLPHDKSMPLLEF 193
182 LRVCHKGNMMLNDRIQNLALQHVLRKAEYIGTLPEPCAFEFERFOELGKMG 241
194 LRAHYKMTMMLNDRISLSALQALRKAEHSLTLOADTPYSEFHHRFOLGKMG 253
242 DTAERLEMIQILLDLLEATDPTLEKFLGRIIPWENVVILTPHGYFAQDNVLAGYPTGG 301
254 DCAKRAQETIHLDLLEAPDPTLEKFLGRIIPWENVVILTPHGYFAQDNVLAGYPTGG 313
302 QVVYILDOVRALLENMLRIKQGGINTPRILITRLIPDAVGTTCGRLKGVTEHSD 361
314 QVVYILDOVRALLENMLRIKQGGINTPRILITRLIPDAVGTTCGRLKGVTEHSD 373
362 ILRVFRTKGIIVRKWISRFKWPYLETYEDVAHEISKELHGTDLIGNKSDGNIVA 421
374 ILRVFRTKGIIVRKWISRFKWPYLETYEDVAHEISKELHGTDLIGNKSDGNIVA 432
422 SLAAHKGVTCTTAHALEKTKYPSDSIYMKLEDKYHSCQPTADLFAMNHTDPIITST 481
433 CLAAHKGVTCTTAHALEKTKYPSDSIYMKLEDKYHSCQPTADLFAMNHTDPIITST 492
482 FOEIAKSDTGVQYESHAFITPGLYRVVHGIDVDFPKFNIVSPGADLSIYFPYTESHR 541
493 FOEIAKSDTGVQYESHAFITPGLYRVVHGIDVDFPKFNIVSPGADLSIYFPYTESHR 552
542 LKHPPEIEDLLYTKVNEEHLCVLNDRNKPIITFMPRLDRVKLTGLVEMCGKNPKLRE 601
553 LKHPPEIEDLLYTKVNEEHLCVLNDRNKPIITFMPRLDRVKLTGLVEMCGKNPKLRE 612
602 LANLVVVGDRKESKDLLEKAEKMKFELIDKYNLNGOFWISSOMNRIRNVELYRITC 661
613 LANLVVVGDRKESKDLLEKAEKMKFELIDKYNLNGOFWISSOMNRIRNVELYRITC 672
662 DTGAFFVQPALYEAFTLVVAMTCGLPTFATCNGSPAETIVHKGSGFNIDPYHGDQAD 721
673 DTGAFFVQPALYEAFTLVVAMTCGLPTFATCNGSPAETIVHKGSGFNIDPYHGDQAD 732

[illegible]

RESULT 15
US-10-425-114-72813

; Sequence 72813, Application US/10425114
; Publication No. US20040034888A1

; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David R.

APPLICANT: Steven E

APPLICANT: JABASKA, JACK E

TITLE OF INVENTION: Nuclear

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21 (53313)B

CURRENT APPLICATION NUMBER: US/10/425.114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEO ID NOS: 73128

SEQ ID NO 72813

LENGTH: 823

TYPE: PRT

ORGANISM: Zee

FEATURE:

OTHER INFORMATION:

US-10-425-114-72813

US-10-425-114-72813

Query Match	77.7%;	Score 3299;	DB 12;	Length 823;
Best Local Similarity	76.9%;	Pred. No. 7.3e-303;		
Matches 619;	Conservative 83;	Mismatches 101;	Indels 2;	Gaps 2

QY	ERALTTHSIRERIDEITLHNRHEIITALLSRIGKSGLOHQQIILEE-EXIPENNNK 61
Db	14 DVLSTRHSYRERIGDSSLNHPNELVAVFTRLKLNKQMLQHOIILAEYNNAIPEAREK 73
QY	62 LANGAFEEVLKASGEAIVLPPWALAVRPAPGVMEYIRVUNVALVVEELTVAEYLHFKEE 121
Db	74 LKOGAFEDVLRAAGEAIVIPPPWALAIRPPGVMEYIRVUNSELAAVEELRAPEYLOFXEQ 13
QY	122 LVDDSGNSGNRYLELDEFPFNSSFFPRLTSSKSINGYEFNRLHSAKLTHDKSNHPLLEF 187
Db	134 LVEEGSPNNNEFLDELDFEPFNASFPSPSLSSKSINGQOFLNRHLSKTLPHDKSNMPLLNF 193
QY	182 LRVCCHKGNMMLNDRIGONLNALOHVLRKAEVLTGLPETERCAEHRFQBIGLERGOM 244
Db	194 LRANHYGNMMLNDRIRSLASALQGRKAEHELTLQADTYSEFHRFQELGIEKMG 255
QY	242 DTAERVLEMIQLLIDLLEATDPCTLEKFLGRIEMVENVILTRPHGYPAOINVLYPDTGG 301
Db	254 DCAKRAQETHILLDLLEADPSTLEKFLGTIRPMVENVILSPHGYPAOINVLYPDTGG 313
QY	302 QVVYIILDOVALENEMELIRIKOGLNITRILITLLELDVAVGTTQGRLEKVVYTHSD 366
Db	314 QVVYIILDOVAMEBEMELIRIKOGLDITRPKILITLLELDVAVGTTQGRLEKVVYTHSD 378
QY	362 ILRVFPRTKGIYAKWISRFEPKWPYLEETTEVVAHEIKELHGTPODILGNXSDGNIVA 421
Db	374 ILRVFPRTENGIVAKWISRF-VPYLETITTDVAHEIKELIGELQANPDILIGNYSDGNIVA 433
QY	422 SLNHLKLVTOCTIAHLEKTKYPPDSDIYWKJLEDKRYHSCQFTADLPAMNHTDEIITST 483
Db	433 CLNHLKMGVTHCTIAHLEKTKYPPNSDLIYWKJFEDHYHSCQFTTDLLAMNHADEIITST 495
QY	482 FOELIAGSKDYTGYESHTAFTLGLRYVNHGIVLPPKKNIVYSPADMEIYTPPYTEGR 54

Dd	493	FOE1AGNDKYQVQESHNAFTWPGLYRVVHGIDVDPPKNIVSPGADLSITFFPYTSHKR	552
Qy	542	LKHPRPEIEDLLYTVNEMENHCYDNDNKRPLTFMPRLDRKYNLTGYEMWGKXNKRE	601
Dd	553	LTSJLHPELLEBLLYSOTENTHEKFVLDNRKPLIFESARIDRKNNLTGLVELGRNRLOE	612
Qy	602	LANLVVVGGDGRBKESKDLEKAEKMKMFELIDKYNLNGQFWMISSQOMNRIRNVELYRYIC	661
Dd	613	LVNLVVGCGDHGNPSPKDEQAEFFKXMFDLIEQYNLNGHIRWISAQMRVRNDELRYIC	672
Qy	662	DTKGFVQPALYLEAGLTVVEMTQGLPRFATCNGRAPIIVHSGSGMINDRPHGQOADD	721
Dd	673	DTKGAFVQPALEAGLTVVEMTQGLPRFATAVGGPARIIVHSGVGHNDIPQSGKASA	732
Qy	722	ILVDFEFEKCKDPSHMDRISQGGKRIIEBKYYTWKTYSEBLTLTVGYGVFWKHVSNLRRR	781
Dd	733	LLVDFDFCQADPSPMSKISQGGQRIIEBKYYTWKLYSEBLMTLTGVGVFWKYVSNLRRR	792
Qy	782	SRRYLEMFYALKYRCLAAEVPAAEE	806
Dd	793	TKRYLEMLYALKYRTMASTVPLAVE	817

Search completed: June 2, 2004, 14:54:51
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 21:13:05 ; Search time 925 Seconds
(without alignments)
7898.058 Million cell updates/sec

Title: US-10-003-405-1
Perfect score: 2625
Sequence: 1 atggcgcagcgtctccac.....tcaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1302.4	49.6	3056	13	BU103683
2	1266.8	48.3	3140	11	AV103630
3	1085.2	41.3	2705	11	AV104856
4	696.6	26.5	1015	14	CD726832

5	677.8	25.8	689	13	BQ406651	BQ406651 GA_EA009
6	677.2	25.8	690	12	BG439960	BG439960 GA_EA000
7	674	25.7	683	13	BQ405173	BQ405173 GA_EA007
8	664.8	25.3	746	14	CD486446	CD486446 CHR2_2P08
9	649.6	24.7	682	13	BQ406766	BQ406766 GA_EA009
10	633	24.1	951	14	CF514627	CF514627 CABU0005
11	630.2	24.0	656	13	BQ409552	BQ409552 GA_EA002
12	623.2	23.7	637	13	BQ415686	BQ415686 GA_EA010
13	623.2	23.7	664	9	AT731292	AT731292 BNGH1910
14	623	23.7	642	12	BG442235	BG442235 GA_EA001
15	620.2	23.6	871	10	BE054845	BE054845 GA_EA003
16	618.6	23.6	633	13	BQ415389	BQ415389 GA_EA009
17	617.4	23.5	653	12	BG443072	BG443072 GA_EA001
18	613.4	23.4	870	14	CD573686	CD573686 UCRPT01_0
19	610	23.2	634	14	BQ415266	BQ415266 GA_EA009
20	608.6	23.2	672	10	AM561929	AM561929 IPRH2001
21	596.2	22.7	945	14	CF213965	CF213965 CCF10081
22	581.8	22.2	881	14	CB893506	CB893506 EST646298
23	580.6	22.1	634	13	BQ407029	BQ407029 GA_EA010
24	577.2	22.0	956	14	CK279779	CK279779 EST725857
25	573.8	21.9	585	13	BQ411883	BQ411883 GA_EA004
26	573.6	21.9	947	14	CK274027	CK274027 EST720105
27	569	21.7	577	9	AT731115	AT731115 BNGH1873
28	568.4	21.7	928	14	CK284157	CK284157 EST746879
29	567.6	21.6	938	14	CK295709	CK295709 EST758423
30	567.2	21.6	688	12	BG445249	BG445249 GA_EA002
31	560.6	21.4	921	14	CK285493	CK285493 EST748215
32	557	21.2	798	14	CD576633	CD576633 UCRPT01_0
33	555	21.1	685	13	BQ406489	BQ406489 GA_EA009
34	553	21.1	999	10	BE16489	BE16489 MG0011.CO
35	550.6	21.0	933	14	CK279264	CK279264 EST725342
36	541.2	20.6	831	14	CF209749	CF209749 CAB20005
37	539.6	20.6	893	14	CK270584	CK270584 EST716662
38	536	20.4	693	14	CD486175	CD486175 CRUS7_1D0
39	535.4	20.4	799	14	CF118999	CF118999 MTU10CS.P
40	533	20.3	798	12	BG583977	BG583977 EST485737
41	530.4	20.2	797	14	CF118943	CF118943 MTU10CS.P
42	526.4	20.1	836	14	CF209223	CF209223 CAB20004
43	526	20.0	806	12	BG581937	BG581937 EST483673
44	525.4	20.0	806	14	CA917502	CA917502 EST641649
45	522.4	19.9	796	14	CB972247	CB972247 CAB10006_

ALIGNMENTS

RESULT 1
BU103683/LOCUS
DEFINITION SCCCR21002G07.g Saccharum officinarum mRNA (Nogueira,F.T.S)
ACCESSION BU103683
VERSION BU103683.1 GI:32815014
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE 1 (bases 1 to 3056)
Nogueira,F.T.S., de Rosa,V.E. Jr., Menossi,M., Ulian,E.C. and Arruda,P.
RNA expression profiles and data mining of sugarcane response to low temperature
Plant Physiol. 132 (4), 1811-1824 (2003)
JOURNAL MEDLINE
PUBMED 12913139
COMMENT Contact: Nogueira FTS
Bioinformatics Lab
Organization for Nucleotide Sequencing and Analysis
C.P. 6176, Campinas, SP 13083-970, Brazil
Tel: 55 19 37881101
Fax: 55 19 37881089

FEATURES
source

Email: tebalidi@unicamp.br.
Location/Qualifiers
1. .3056
/organism="Saccharum officinarum"
/mol_type="rRNA"
/db_xref="taxon:4547"
/clone_lib="Saccharum officinarum mRNA (Nogueira, F.T.S)"

ORIGIN

Query Match 49.6%; Score 1302.4; DB 13; Length 3056;
Best Local Similarity 71.3%; Pred. No. 2.2e-234;
Matches 1746; Conservative 0; Mismatches 697; Indels 6; Gaps 2;

7 GAGCGTCTCACTCGGCTCCAGTCTCCGTGACGCTTGGATGAGACCTTCTTCT 66
Db GACCGTCTGAGCGGCTCCAGCGCTGAGGAGCGCATTTGGGATTCACCTCTCC 2805
67 CACAGAACGAGATTTTGGCCCTGCTCCAGAGATCGAGGCGAAAGAAAGAAATTC 126
Db CACCCCAATGAGCTTGTGCGCTTTCACAGGCTGAAACCTTGGAAAGGATATCTG 2745
127 CAACACATCAATATTCTAGAGTTTGA--AGCTATCCCTGAGAGAGACGAAAGAG 183
Db CAGCCCAACAGATATTGCGAGTACACAGTGCATTCCTGAGGCTGAGCGTGAAG 2685
184 CTCGCTAATGATGATTTTGAAGATTTGAAGGCTAGTCAGAAAGCATGCTTCT 243
Db TTGAAGATGAGCGCTTGAAGATGCTGAGGAGCTCAGAAAGCAATTTATCCCC 2625
244 CCATGGGTGACATGCTGCTGCTCCAGAGCGCTGCTTGGAGATCATTTAGATGAT 303
Db CCATGGGTGACATGCTGCTCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2624
304 GTTCAGCGCTTGTGTTGAGAGATCTCATGTTGCTGATCTCCATTCAGAGAGAG 363
Db GTTCAGAGCTGCTGCTGTTGAGAGCTGAGAGTCCAGAGTACCTGATTCAGAGAG 2505
364 CTGTGTGAGAGATTCATTAATGAACTTTGTTTGAATTGATTTTGAAGCCCTTCA 423
Db CTGTGTGAGAGAGGCCCCCAACAACTTTGTTTGAAGCTGATTCAGATTCAT 2504
424 TCATCATTCCTCCCGCCCACTTTTCAAAATCCATTTGATATGATGATGATCTTAAT 483
Db GCATTCCTCCCGCTTCTTCTGCAAAATGCTATGCAATGATGATGATGATCTTCAAC 2444
484 GGTCACTTTGCGCAAAATTTGTCATGACAAAGAGAGATGACCTTTGCTCGAATTC 543
Db AGGCACTGTATCAAACTCTTCCATGACAAAGAGAGATGATACCTTCTCAATTC 2384
544 CTCAAGTTCATTTGTCAGAGGCGAAACATGATGTTGAATGACAGAAATTCAGAACTTG 603
Db CTTCGCGCCCAACATCAAGGAGTACATGATGTTGAATGACAGAAATTCAGAACTTG 2324
604 AATGCTCTTCAACATGTTTGAAGAAAGCAGAGAGATCTTTGTAACCTTCTCTGAG 663
Db AGTGTCTGCAAGGCGCTCTGAGAAAGGCTGAGAGCACTGTCCACCTTCCAGCTAT 2264
664 ACACATGTCGCAATTCGAAACACCGGTTCCAGAAATCGGTTTGAAGAGTGGGT 723
Db ACCCAATCTGTAATTTCAACAGATTCAGAGAACTTGGTCTGAGAAAGGTTGGGT 2204
724 GACACCGAGAAACGCGTCTCGAGATGATCAATCTCTTTTGAATCTTTTGAAGCACT 783
Db GATCGCGCTAAGCGCGCGCAGAGACTATTCACCTCTTCTGACCTTCTGAGGCCCA 2144
784 GATCCTTGACACCTTGAGAGATTTCTTGGAGAAATCCCAATGCTTAAATGTTGAT 843
Db GATCGCTGACACCTTGAGAGATTTCTTGGAGATTCCTCATGATGTTTAAATGTTATC 2084
844 CTCATCTCCCAAGATATCTGCTCAAGACATGTTTGGGATATCCGACACCGTGGG 903
Db CTCATCTCCCAAGATATCTGCTCAAGATGCTTGGGATATGCTTGGGATATCCGACAC 2024
2024 CTCATCTCCCAAGATATCTGCTCAAGATGCTTGGGATATGCTTGGGATATCCGACAC 1965

904 CAGGTGTTTACATCTTTCGATCAAGTCCGAGCTTTGAGAAATGAGATGCTCCGATTA 963
Db CAGGTGTTTACATCTTTCGATCAAGTCCGAGCTTTGAGAAATGAGATGCTCCGATTA 1964
964 AAGCAACAGGATCAATCAACCCCTGAAATCTCATTTACTTACAGCTTCTCTGAT 1023
Db AAGCAATGCTGCTTTCGATCAATCAACCCCTGAAATCTCATTTACTTACAGGTTCTCTGAT 1904
1024 GCTGTCGAAACCAATGTCGTCACGATCTGAGAAATGATATACGAAACGAGCACTCGAT 1083
Db GCAATGTCGAAACCAATGTCGTCACGATCTGAGAAATGATATACGAAACGAGCACTCGAT 1844
1084 ATTTCTGAGTACCTTTCAGAACGAAAGGAAATTTGTTGAAATGAGATCTCAAGATT 1143
Db ATTTCTGAGTACCTTTCAGAACGAAAGGAAATTTGTTGAAATGAGATCTCAAGATT 1784
1144 GAAAGTCTGAGCAATCTTGGAAACCTTACAGAGATGTTGCTCATGAAATCTCCAA 1203
Db G--AAGTCTGAGCAATCTTGGAAACCTTACAGAGATGTTGCTCATGAAATCTCCAA 1724
1204 GAGTTGACGCGACGCGCATCTGATCATCGAAACCAACGCGCATATGCTGCC 1263
Db GAGTTGACGCGCATCTGATCATCGAAACCAACGCGCATATGCTGCC 1667
1264 TCTTCTGCGACATTAATTTAGTGTCAACAGTGCACATCGCCCATGCTTTGAGAG 1323
Db TGTGTTGCTGCGACATTAATTTAGTGTCAACAGTGCACATCGCCCATGCTTTGAGAG 1607
1324 ACAAAATATCCAGATTCAGATATCTATTTGAGAAAGCTGAAAGCAATATCATTTCT 1383
Db ACTAATATCCAGATTCAGATATCTATTTGAGAAAGCTGAAAGCAATATCATTTCT 1547
1384 TGCCAAATTTACAGCTGATCTTTTTCATATGACATACAGATTTTATCATCATCACT 1443
Db TGCCAAATTTACAGCTGATCTTTTTCATATGACATACAGATTTTATCATCATCACT 1487
1444 TTCCAGAAATTTGAGAGAGAGAGACATGTTGTCATATGAGAGGCAATGCTTTT 1503
Db TTCCAGAAATTTGAGAGAGAGAGACATGTTGTCATATGAGAGGCAATGCTTTT 1427
1504 ACTCTCTGCTGCTCTACCGGTTGTCATGATGTCATGATGTCATGATGTCATGATGTCAT 1563
Db ACTCTCTGCTGCTCTACCGGTTGTCATGATGTCATGATGTCATGATGTCATGATGTCAT 1367
1564 ATTGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACACGAGAGAGAGAG 1623
Db ATTGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACACGAGAGAGAGAG 1307
1624 TTGAAGATTTTCCATCTGAGATGAGAACTTTCTTTACACAAAGTTGAGATGAGAA 1683
Db TTGAAGATTTTCCATCTGAGATGAGAACTTTCTTTACACAAAGTTGAGATGAGAA 1247
1684 CACTTATGTCGTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
Db CACTTATGTCGTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187
1744 GGTGTCAAGAACTTTAACCGGATCTGTCAGAGTGTGCGGCAAGAACCAAGTTGCTG 1803
Db GGTGTCAAGAACTTTAACCGGATCTGTCAGAGTGTGCGGCAAGAACCAAGTTGCTG 1127
1804 TTGCTTACCTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1863
Db TTGCTTACCTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1067
1864 AAGCTGAAATGAGAGAAATGTTGAGTGTGACAGATGATGATGATGATGATGATGAT 1923
Db AAGCTGAAATGAGAGAAATGTTGAGTGTGACAGATGATGATGATGATGATGATGAT 1007
1924 AGATGATATCATCTCAATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAG 1983
Db AGATGATATCTGAGATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 947
1984 GACAGAAAGGTCCTTTTGTACAGCTGATTTGATGAGAGCTTTGATGAGAGCTTGTG 2043

QY 961 ATAAAGCAAGACATCAATCAATCAATCCCTGAAATCTCTATTTACTATGACCTTCTCT 1020
 DB 1138 ATCAAGAGCTGTGGTCTGATCATCAAGCCGAAGATCTTATTTGACCAAGGTTGCTCTCT 1197
 QY 1021 GATGCTGCGAAGCAATG-CGGTCAAGCACTTGAGAAAGATATACGGAACAGACACTC 1079
 DB 1198 GATGCACTGGACACCTGTGGGCGACAGGCTTTAGAAAGGTCCTTGGACACGACACTG 1257
 QY 1080 GGATATCTTGCAGATCCCTTCAGAAAGAAAGGAAATGTTGCAAAATGATCTCAAG 1139
 DB 1258 CCAATCTTGGCGTGCATTCAGAAAGAAAGGAAATCGTTGCAAGTGGATCTCGCG 1317
 QY 1140 ATTGAAAGAGCTGTGGCAATCTTGAAACCTCAACAGAGATGTTGCTCATGAATCTC 1199
 DB 1318 ATTTG---AAGTGTGGCCGTACCTGTAGACTTACATGATGACGTGGCGCAATGAGATTC 1374
 QY 1200 CAAGAAGTTGACGGCAGCGCAATCTGATCATCGGAAACNACAGCGGCAATATGCT 1259
 DB 1375 TGAAGAGCTTCAAGGCAATCTGACCTGATCATCGGAAACTACAGTGAAGAACTTGT 1434
 QY 1360 CGCCTCTTGCCTGACATTAATTTAGTGTCAACAGTGCACATCGCCCATGCTTTGGA 1319
 DB 1435 TGGGTGTTGCTCGGCCCAAGATGGGTGTTACTCACTGTACATGCGCATGCGCTTGA 1494
 QY 1320 GAAGCAAAATATCCAGATTCAGATATCTATTGGAAGAGCTTGAAACAAATACATTT 1379
 DB 1495 GAAATCTAAGTACCTTAATCCGACCTCTCACTGAAAGAGTTTGAAGATCACTACACTT 1554
 QY 1380 CTCTTGCATTTACAGTGAATCTTTTGTGAAGTGAACATACAGTTTCAATCATCAACAG 1439
 DB 1555 CTCGTGCAAGTTCACACTGATCTTATTTGCAATGAACATGCGACTTCAATCATCAACAG 1614
 QY 1440 TACTTTTCAGAAATTTGAGAGAGCAAGCACTTTGTGCAATTCAGAGACACATGTC 1439
 DB 1615 TACCTTCCAAAGATTCGCGGAAACAGAGACACGTCGCGCACTGACGTCAACATGTC 1674
 QY 1500 TTTGACTCTCTGCTCTCAACCGTGTGTGACATGATTCGATGTTTATTCCTCAATTT 1559
 DB 1675 GTTCAATGCTGCTGCTGCTGACCGGTGTGCTCAACGCGCATTTATGTTGTGACCCCAATTT 1734
 QY 1560 CAACATTTTCTCCCTGCTGTGATGATGAGATATCTTCCCTTACACCGAAGAGAACCG 1619
 DB 1735 CAACATCTGTCTCTGCGGCGGACCTGTCACTTACTTCCCGTACACCGAGTGCACAA 1794
 QY 1620 GAGGTGGAAGATTTCCATCTCGATGAGAGACCTTCTTTTACCAAGTTGAGATGA 1679
 DB 1795 GAGGTGACCTCCCTTCAACCGGAGATGAGAGCTCTCTGACAGCCAAACCGAGAAC 1854
 QY 1680 AGAACATTTATGTGTCTCAATGACCGCAACAGCCAAATCTGTTCAATGCGCAAGCT 1739
 DB 1855 GAGACCAAGTTCTGTTGAAAGACAGAGAACAGCCAAATCTTCTCATGCTGCTCT 1914
 QY 1740 TGATGTTTCAAGAACTTAAACCGGACTGTGAGAGTGTGGCGCAAGAACCCAAAGTTGCG 1799
 DB 1915 CGACCGTGTGAAGAACTTGACCTGAGTGTGAGAGTGTGACGCGGAAACAGCGCTGCA 1974
 QY 1800 TGAATGCTAACTCTGATGTTGTAGTGTGATGAGGAAAGGATCTTAAAGATTTGGA 1859
 DB 1975 GAGGTGTGAACTCTGTGTGTGTGTGCGGCAACATGGAACCTTCCAGAGCAAGGA 2034
 QY 1860 AGAGAGGCTGAATGAAGAAATGTTGAGCTGATGCAAGATCAACTTGAACGCGCA 1919
 DB 2035 GAGAGAGGCGAGTTCAAGAAAGATGTTGACCTTCAAGCAGTCACTGAACGCGCA 2094
 QY 1920 ATTCAATGATATATCTCAAAATGAACAGAAATCGGAAATGTTGAATTTTACCTTACAT 1979
 DB 2095 CATCGCTGAGATCTCGCCAGATGAACCGGCTCGCAACGCGAGCTGTACCTGTATAT 2154
 QY 1980 TTGGAACAGAAAGGCTTGTGTACAGCTGCAATTTGATGAAGCTTTGATTTGACAGT 2039
 DB 2155 CTGCGACACCAAGGCGCTTGTGTGACAGCTGCTTTCTTACAGAGCTTTGCGGCTGACGT 2214
 QY 2040 TGTGAGGCAATGACTTGGGTTTGCACAACTTGGCAACGTTGAACGTTGACCAAGCGCA 2099

DB 2215 GGTGAGGCGCATGACCTGCGGCTGCGCCAGCTTTGCCACAGCTTACGCGGCTCGGCGCA 2274
 QY 2100 GATTATTTGTCATGGGAATCTGGTTTCAACATTTATCTTACCATGCTGATCAAGCTGC 2159
 DB 2275 GATCATGCTGACCGCGGTGTGTGCTTACATGACCTTTACAGAGGCGCAAGGCGCTC 2334
 QY 2160 TGACATCTGCTGATTTCTTTGAAAGATGTGAAGAAAGATCATCTCACTGGGATAGAT 2219
 DB 2335 GGCCCTGCTGTGACCTTCTTCAAGATGCGACGCGGACCGGACCTAGACCAAGAT 2394
 QY 2220 CTCCCAAGAGGCTTGAAGCAATGAGAGAAATATATCATGAAAGATTTACTGGAGAG 2279
 DB 2395 CTCCAGGCGGCTTCCAGGCTATCGAGAGAAATACCTGGAAGCTTACTGGAGAG 2454
 QY 2280 ACTATTAACCTTGAACAGAGATGATGATTTCTGGAAGCATTTTCAACTTGAACGCG 2339
 DB 2455 GCTATGACCTTACACGCGCGGTGACGCGTTCTGGAAGTACTGTCAACCTGAGAGCG 2514
 QY 2340 TGAGAGTGTGCTTACCTTGAATGTTTATGCTTCTTAAGTACCGTAAGCTGCTGAATC 2399
 DB 2515 CGAGACCGGCGGTACCTGAGATGTTGACGCGCTCAAGTACCGCACATGCGGAGAC 2574
 QY 2400 AGTTCCATTTGCGACAGAG 2418
 DB 2575 CGTCCGCTGCGCGGTGAG 2593

RESULT 3
 AY104856
 LOCUS 2705 bp mRNA linear HTC 16-Oct-2002
 DEFINITION Zea mays PC0079420 mRNA sequence.
 ACCESSION AY104856
 VERSION AY104856.1 GI:21207934
 KEYWORDS HTC.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 2705)
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitelc, M.S.,
 Arthur, L.W., Hanafey, M., Morzante, M. and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 2705)
 Coe, E.H.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.
 location/Qualifiers
 1..2705
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:634420"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont configs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Query Match 41.3% Score 1085.2; DB 11; Length 2705;
 Best Local Similarity 66.5%; Pred. No. 1.2e-193;
 Matches 1587; Conservative 0; Mismatches 794; Indels 6; Gaps 2;

```

Qy 38 GTGACGTTTGATGAGACCTTCTTGCTCAGAGAGAGATTTGGCTTGCTCTCA 97
Db 1 GCGACCGGCTGAGAGACCTTCCACGCGCACCGCAAGAGCTGTGCGCTCTGCTCA 60
Qy 98 GGATGAGGCGAAGAGAGAGATTTGCAACACCATCAATTTATTTAGATTGAG 157
Db 61 AGTACGTAAACAGAGAGAGAGAGATTTGCAACACCATCAATTTATTTAGATTGAG 120
Qy 158 CTATCCCTGAGAGAGAGAGAGAGATTTGCAATTTGATTTGAGATTGAGAG 217
Db 121 AGGTCCAGAGAGAGAGAGAGAGATTTGCAATTTGATTTGAGATTGAGAG 180
Qy 218 CTAGTCAGAGAGAGAGAGAGAGATTTGCAATTTGATTTGAGATTGAGAG 277
Db 181 CCGGCGAGAGAGAGAGAGAGAGATTTGCAATTTGATTTGAGATTGAGAG 240
Qy 278 GTGTTGGAGAGATTTGATGAGATTTGCAACGCTTGTTGAGAGATTTGAG 337
Db 241 GAGTTGGAGAGATTTGATGAGATTTGCAACGCTTGTTGAGAGATTTGAG 300
Qy 338 CTGAGATCTCCATTTCAAGAGAGAGATTTGATGAGAGATTTGATGAGAG 397
Db 301 CGAGATCTCCATTTCAAGAGAGAGATTTGATGAGAGATTTGATGAGAG 360
Qy 398 TGAATGAGATTTGAGAGAGAGATTTGATGAGAGATTTGATGAGAGAG 457
Db 361 TCGAGATTTGAGAGAGAGAGATTTGATGAGAGATTTGATGAGAGAG 420
Qy 458 TTGTAATGAGAGAGATTTGATGAGAGATTTGATGAGAGATTTGATGAG 517
Db 421 TTGTAATGAGAGAGATTTGATGAGAGATTTGATGAGAGATTTGATGAG 480
Qy 518 AGAGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 577
Db 481 ATGTAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 540
Qy 578 TGTGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 637
Db 541 TGTGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 600
Qy 638 AGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 697
Db 601 AGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 660
Qy 698 AAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 757
Db 661 AGTGGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 720
Qy 758 TCTTTGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 817
Db 721 TCTTTGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 780
Qy 818 TCCCATGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 877
Db 781 TCCCATGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 840
Qy 878 TTTTGGGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 937
Db 841 TATTAGGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 900
Qy 938 TGAATGAGAGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGAGAG 997
Db 901 TAGAATGAGAGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGAGAG 960
Qy 998 TCAATGAGAGAGATTTGAGAGAGAGATTTGATGAGAGAGATTTGAGAG 1057
Db 961 TCAATGAGAGAGATTTGAGAGAGAGATTTGAGAGAGATTTGAGAGAG 1020

```

```

Qy 1058 AAGTATGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1117
Db 1021 GAAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1080
Qy 1118 TTGTTGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1177
Db 1081 TACTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1137
Qy 1178 AGAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1237
Db 1138 AGAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1197
Qy 1238 ACNACGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1297
Db 1198 ACTACGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1257
Qy 1298 GCACATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1357
Db 1258 GCACATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1317
Qy 1358 AGCTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1417
Db 1318 ATTCGATGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1377
Qy 1418 ATACGATTTGATGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1477
Db 1378 ATGCTGATTTGATGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1437
Qy 1478 GTCAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1537
Db 1438 GACGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1497
Qy 1538 TCGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1597
Db 1498 TCGATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1557
Qy 1598 TCCCTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1657
Db 1558 TTCCATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1617
Qy 1658 TTTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1717
Db 1618 TTTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1677
Qy 1718 TTCTGATGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1777
Db 1678 CCTCTTCTGATGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1737
Qy 1778 GCGGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1834
Db 1738 TTCTGATGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1797
Qy 1835 GCGGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1894
Db 1798 ATGTCAGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1857
Qy 1895 TCGAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1954
Db 1858 TCGAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1917
Qy 1955 GAAATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2014
Db 1918 GTACGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 1977
Qy 2015 TGTATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2074
Db 1978 TGTATGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2037
Qy 2075 CAACGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2134
Db 2038 CGAGGAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2097
Qy 2135 ATCTTAAAGAGAGAGAGAGAGATTTGATGAGAGAGATTTGATGAGAG 2194

```

Db 2098 ACCGTCACCCGACAGAGCTGTATCTGATGCGGACCTCTTCGACCGGTGCAAGC 2157
Qy 2195 AAGATCCATCTCACTGGGATAGATCTCCCAAGAGGCTTGAAAAGATAGAGAGAGT 2254
Db 2158 AAGACCCAGATCACTGGGTGAAATATCTGAGAGCAGGGCTCAGGCGATATACGAGAGT 2217
Qy 2255 ATACATGGAAGATTATCTCGAGAGAGACTATTTGACCTTGACAGAGGTATGAGATTCTGGA 2314
Db 2218 ACACATGGAAGATTATCTCAGAGAGGTTGATGACACTGGCGGGGTCTACGGTTTCTGGA 2277
Qy 2315 AGCATGTTTCCAACTTGAAGCGCTGAGTCCGTCTTACCTTGAGATTGTTATGCTC 2374
Db 2278 AGTACGTGTCAAGTTCAGAGAGCGGAGACGAGCGCTACCTTGAGATTCTACATAC 2337
Qy 2375 TTAAGTACCTGAGCTGCTGATGATCAGTTCCATTGGCAGAGATGA 2421
Db 2338 TGAAGTCCGCGAGCTGCGCAAGACCGTGCCTTGCAATTGACCA 2384

RESULT 4
CD726832/c 1015 bp mRNA linear EST 26-JUN-2003
LOCUS CD726832
DEFINITION Cucurbita pepo testa subcloned cDNA Cucurbita pepo cDNA
clone SUST similar to Sucrose synthase, mRNA sequence.
ACCESSION CD726832.1 GI:32277679
VERSION
KEYWORDS
SOURCE
ORGANISM
Cucurbita pepo
Cucurbita pepo
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.

REFERENCE
1 (bases 1 to 1015)
Bezold, T.N., Mathews, D., Loy, J.B. and Minocha, S.C.
Molecular analysis of the hull-less seed trait in pumpkin:
Expression profiles of cell wall related genes during development
Unpublished (2003)
CONTACT: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold
JOURNAL
COMMENT

Dr. Minocha
University of New Hampshire
Rudman Hall, Durham, NH 03824, USA
Tel: 603 862 3840
Fax: 603 862 3784
Email: sminocha@cisunix.unh.edu
Degenerate primers and Tag were used to amplify cDNA for TOPO TA
(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three
times using the Dyanamic ET Terminator Sequencing kit (Amersham
Pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified
by NCBI BLAST(X).

FEATURES

source
1..1015
Location/Qualifiers
/organism="Cucurbita pepo"
/mol_type="mRNA"
/db_xref="taxon:3663"
/clone="SUST"
/dev_stage="20 days post-anthesis"
/clone_11b="Cucurbita pepo testa subcloned cDNA"
/note="Organ: Testa; Total RNA was isolated from 20 day
post-anthesis testa tissue and used in a subtraction
hybridization procedure as according to the Clontech
PCR-Select cDNA Subtraction kit. (PT1117-1) (Clontech, Palo
Alto, CA)."

ORIGIN

Query Match 26.5%; Score 696.6; DB 14; Length 1015;
Best Local Similarity 80.8%; Pred.No.1.3e-120;
Matches 823; Conservative 1; Mismatches 191; Indels 3; Gaps 1;
Qy 563 AGGCAAGACATGATGTTGATGACAGAAATTCAGAACTTGAAATCTCTCAACATGTTT 622
Db 1015 AGGCAAGACATGATGTTGATGATGATGATTCGAACTTGAAATCTCTCAACATGTC 956

Qy 623 TGAGAAAGCAGAGAGATATCTTGTACCTTACCTCTGAGACACATATGTCGGAATTGC 682
Db 955 TGAGAAAGCAGAGAGATATCTTGTACCTTACCTCTGAGACACATATGTCGGAATTGC 896
Qy 683 AAGACCGGTTCCAGGAATCGGTTTGGAAAGAGGTTGGGGGTGACACCGCAGAAAGCGTGC 742
Db 895 CGCTGAAGTTCCAGGAATGCTTGGCTCGAGAGAGGTTGGGGGTGACACCGCAGAGTGTAC 836
Qy 743 TCGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCACTGATCTTTGACCCCTTGAGA 802
Db 835 TTGAATGATCCAACTCTTTTGGATCTTCTTGAAGCACTGATCTTTGACCCCTTGAGA 776
Qy 803 AGTTCCTTGGAGAAATCCCATGCTTTCATATGTTGTATCTCACTCTCCACGATATCT 862
Db 775 AGTTCCTTGGAGAAATCCCATGCTTTCATATGTTGTATCTCACTCTCCACGATATCT 716
Qy 863 TCCTCAAGCAATGTTTGGGGGTATCCCGACACCGGTGGCAGGTTTACATCTTGG 922
Db 715 TTGACAAAGATACGTCCTTGGGTTACCCCGACACCGGTGGCAGGTTTACATCTTGG 656
Qy 923 ATCAAGTCCGAGCTTTGAGAAATGATGCTCTCCGTATTAAGCAACAGACTCAACA 982
Db 655 ACCAAGTTGCTGTTGGAGATGAATGCTTCAACGATATTAACACCAAGTATTGACACA 596
Qy 983 TCACCCCTCGAAATCTCTATTTATCTAGACTCTTCTCGATGCTGTGGAAACAATGCG 1042
Db 595 TAACTCTCTGATTTCTATTAATCAACAGACTCTCTCCAGAGATGAGAAAGCACTGCA 536
Qy 1043 GTCAAGCACTGAGAAAGTATACGGAACAGAGCACTCGATATTTCTGAGTACCTTCA 1102
Db 535 CACAAGCTTCGAAAGGATATACGGAACAGAGCACTCGATATTTCTGAGTACCTTCA 476
Qy 1103 GAACAGAAAGGAAATGTTTGAATGATCTCAAGATTGAAAGTCTGGCCATATCT 1162
Db 475 GAAATGAGAAAGGATGCTCGAAAGTATCTCAAGATTG---AAGTCTGGCCTTAC 419
Qy 1163 TGGAACTTACACAGAGATGTTGCTATGAAATCTCCAAAGATGTCAGCGACGCCAG 1222
Db 418 TAGAGACTTACACAGAGATGTTGACAGAACTTCCAAAGATGTCAGCGAAAGGCCG 359
Qy 1223 ATCTGATCATGGAAGAAACAGAGCGAGCAATATGTCGCTCTGCTGACATTAAT 1282
Db 358 ATCTGATCATGGAAGAAACAGAGCGAGCAATATGTCGCTCTGCTGACATTAAT 299
Qy 1283 TAGGTGTCAACAGTGAACCATGCGCCATGCTTTGAGAAAGCAAAATATTCAGATTGAG 1342
Db 298 TAGAGTTTACACAGTATGTTGCTCAACGCTTGGAGAAACCAAAATATTCAGATTGAG 239
Qy 1343 ATATCTATGGAAGAAAGCTTGAAGACAAATATCACTTCTCTTGGCAATTTACAGTATC 1402
Db 238 ATATCTATGGAAGAAATTTGAGAGATTAATATCACTTCTCAATGCAAGCTGATTC 179
Qy 1403 TTTTGGCAATGAACATATACAGATTTCATCATCAACAGATCTTCCAGAAATTTGAGAGAA 1462
Db 178 TTATTTGCAATGAACATATACAGATTTCATCATCAAGATCTTCCAGAAATTTGAGAGAA 119
Qy 1463 GCAAGACACTGTTGTGATATACAGAGCAACACTGCTTTGACTCTTCTGCTGATAC 1522
Db 118 GCAAGATATGTTGCAATATACAGAGTCACTGCTTCACTCTTCCGAGTCTTAC 59
Qy 1523 GTGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1580
Db 58 GGTGCTGTCACGGAATGACGCTGTTGATTCCTCAAAATTCMACTGTTGACCTGAGC 1

RESULT 5
BQ406651 689 bp mRNA linear EST 22-MAY-2002
LOCUS BQ406651
DEFINITION GA_Ed0097B05f Goseypium arboresum 7-10 dpa fiber library Goseypium
arboresum cDNA clone GA_Ed0097B05f, mRNA sequence.
ACCESSION BQ406651
VERSION BQ406651.1 GI:21094338
KEYWORDS EST.

SOURCE	ORGANISM
Gossypium arboreum	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE AUTHORS	Ming, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A. An integrated analysis of the genetics, development, and evolution of the cotton fiber
TITLE	Unpublished (2006)
JOURNAL COMMENT	Contact: Ming RA Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwm@clermson.edu Total High Quality bases = 565 Seq primer: TAATCGACTCACTTAAAGG High quality sequence start: 2 High quality sequence stop: 677. Location/Qualifiers
FEATURES	1..689 /organism="Gossypium arboreum" /mol_type="mRNA" /strain="AKA" /cultivar="8400" /db_xref="taxon:29729" /clone="GA_Ed0097B05F" /cruise_type="Fibers isolated from bolls harvested 7-10 dpa" /lab_host="H. coli" /clone_lib="Gossypium arboreum 7-10 dpa fiber library" /note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
ORIGIN	
Query Match	25.8%; Score 677.8; DB 13; Length 689;
Best Local Similarity	99.0%; Pred. No. 5,2e-117; Indels 0; Gaps 0;
Matches	682; Conservative 0; Mismatches 7;
Db	1644 GATCGAAGACCTTTCTTTACACCAAAGTTGAGATGTAAGAACAATTATGTGTCATATGA 17033 1 GATCGAAGACCTTTCTTTACACCAAAGTTGAGATGTAAGAACAATTATGTGTCATATGA 60
Oy	1704 CCGCAACAAGCCAAATTCCTGTTCAACAATGCCAAGCTTGATGCTGCACGAACCTTAACCGG 17635
Db	61 CCGCAACAAGCCAAATTCCTGTTCAACAATGCCAAGCTTGATGCTGCACGAACCTTAACCGG 120
Oy	1764 ACTGTCGAGTGTGTCGGCGCAAGAACCACAAAGTGTGCTGATGTGCTAACCTGTAAGTTGT 18233
Db	121 ACTGTCGAGTGTGTCGGCGCAAGAACCACAAAGTGTGCTGATGTGCTAACCTGTAAGTTGT 180
Oy	1824 AGGTGTGATVAGCGCAAGAAAGAACTTAAGATTTGGAAGAGAAAGCTGAATGAAGAAAT 18933
Db	181 AGGTGTGATVAGCGCAAGAAAGAACTTAAGATTTGGAAGAGAAAGCTGAATGAAGAAAT 240
Oy	1884 GTTTGAGCTGATTCGACAAAGTCAACTTGAACGGCCCAATTCAGATGATATCATCTCAAT 19438
Db	241 GTTTGAGCTGATTCGACAAAGTCAACTTGAACGGCCCAATTCAGATGATATCATCTCAAT 300
Oy	1944 GAACAGAAATCCGAATAATGTTGAACCTTACCGATATTTGCGACACGAAGGTGCTTTGT 20033
Db	301 GAACAGAAATCCGAATAATGTTGAACCTTACCGATATTTGCGACACGAAGGTGCTTTGT 360
Oy	2004 ACAGCTTGATTTGATGAAGCTTTGATTTGACAGTTGTGGAAGGCAATGACTTGGCGTTTT 20633
Db	361 ACAGCTTGATTTGATGAAGCTTTGATTTGACAGTTGTGGAAGGCAATGACTTGGCGTTTT 420
Oy	2064 GCCAATTTGCGACACTGTAAACGATGACACAGCGAGATTAATGTCATGGAATAATCTGG 21233
Db	421 GCCAATTTGCGACACTGTAAACGATGACACAGCGAGATTAATGTCATGGAATAATCTGG 480

QY	2124	TTTCAACATTGATCCTTACCATGGTATGATCAAGCTGCGACATCACTCGTCGATTTCTTTGA	2183
Db	481	TTTCAACATTGATCCTTACCATGGTATGATCAAGCTGCGACATCACTCGTCGATTTCTTTGA	540
QY	2184	AAAGCTTAGAAGATCCATCTCACTGGATAGATCTCCCAAGAGGCTTGAACGAT	2243
Db	541	AAAGCTTAGAAGATCCATCTCACTGGATAGATCTCCCAAGAGGCTTGAACGAT	600
QY	2244	AGAGGAGATATACATGGAAGATTTACTCGAGAGACTATTTGACCTTGACGAGAGTGA	2303
Db	601	AGAGGAGATATACATGGAAGATTTACTCGAGAGACTATTTGACCTTGACGAGAGTGA	660
QY	2304	TGGAATTCGAGAGCATGTTTCCAACTTG	2332
Db	661	TGGAATTCGAGAGCATGTTTCCAACTTG	689
RESULT 6	GC439960	690 bp	mRNA
LOCUS	GC439960	690 bp	mRNA
DEFINITION	GA_Ea0005G07f Gossypium arboreum 7-10 dpa fiber library Gossypium		
ACCESSION	arboresum CDNA clone GA_Ea0005G07f, mRNA sequence.		
VERSION	GC439960		
KEYWORDS	GC439960.1 GI:13349610		
SOURCE	EST.		
ORGANISM	Gossypium arboreum		
REFERENCE	Gossypium arboreum		
AUTHORS	Eukaryote, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosoids; euroside II; Malvales; Malvaceae; Malvoideae; Gossypium.		
TITLE	1 (bases 1 to 690)		
JOURNAL	Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.		
COMMENT	An integrated analysis of the genetics, development, and evolution of the cotton fiber unpublished (2000)		
CONTACT	Contact: Wing RA		
INSTITUTE	Clemson University Genomics Institute		
ADDRESS	Clemson University		
TELEPHONE	100 Jordan Hall, Clemson, SC 29634, USA		
TELEPHONE	Tel: 864 656 7288		
TELEPHONE	Fax: 864 656 4293		
EMAIL	Email: rtwing@clemson.edu		
SEQ PRIMER	Seq primer: TTAATGACACTACTATAGCG		
High quality sequence stop: 684.			
Location/Qualifiers			
1. 690			
/organism="Gossypium arboreum"			
/mol_type="mRNA"			
/strain="AKA"			
/cultivar="8400"			
/db_xref="taxon:29729"			
/clone="GA_Ea0005G07f"			
/issue_type="Fibers isolated from bolls harvested 7-10 dpa"			
/lab_host="E. coli"			
/clone_id="Gossypium arboreum 7-10 dpa fiber library"			
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"			
ORIGIN			
Query Match	25.8%	Score 677.2;	DB 12; Length 690;
Best Local Similarity	98.8%;	Pred.No. 6.7e-117;	
Matches 682; Conservative	0;	Mismatches 8;	Indels 0; Gaps 0;
QY	1326	AAATATCCGATTCAGATATCTTATGGAAGACCTTGAAGACAAATACATTTCTCTTG	1385
Db	1	AAATATTCGATTCAGATATCTTATGGAAGACCTTGAAGACAAATACATTTCTCTTG	60
QY	1386	CCAAATTACAGCTATCTTTTGGCAATGAACCAATACAGATTTTCATTCACCAAGTACTTT	1445
Db	61	CCAAATTACAGCTATCTTTTGGCAATGAACCAATACAGATTTTCATTCACCAAGTACTTT	120
QY	1446	CCAGGAATTGCGAAGACGAAGACACTGTTGGTCAATACGAGAGCCACACTGCTTTTAC	1505

Db	Accession	Version	Keywords	Source	Organism	Reference Authors	Title	Journal	Comment
Db	121	CCAGAAATTGCAGGAAGCAAGACACGTGTGTCTCAATGATATCCACATCTGCTTTAC							
Qy	1506	TCCTTCCTGCTCTACCGGTGTGTGACATGATCGATGCTTTGAATCCAAATTCACAT							
Db	181	TCCTTCCTGCTCTACCGGTGTGTGACATGATCGATGCTTTGAATCCAAATTCACAT							
Qy	1566	TGTTTCCCTGCTGCTGATATGAGATATTACTTCCCTTACACCGAAGGAAGCGGAGTT							
Db	241	TGTTTCCCTGCTGCTGATATGAGATATTACTTCCCTTACACCGAAGGAAGCGGAGTT							
Qy	1626	GAAGCATTTCCATCTCTGAGATCGAAGACCTTCTTACACCAAGTTGAGATGAAGAACA							
Db	301	GAAGCATTTCCATCTCTGAGATCGAAGACCTTCTTACACCAAGTTGAGATGAAGAACA							
Qy	1686	CTTATGTGTGCTCAATGACCGCAACCAATTCCTTACCAATGCTTGAATCG							
Db	361	CTTATGTGTGCTCAATGACCGCAACCAATTCCTTACCAATGCTTGAATCG							
Qy	1746	TGTCAAGAACTTAAACCGGATCTGTGAGATGTGGGCAAAACCAAGTTGCTGAGTT							
Db	421	TGTCAAGAACTTAAACCGGATCTGTGAGATGTGGGCAAAACCAAGTTGCTGAGTT							
Qy	1806	GGCTTAACCTGCTAGTTGATGAGTGGATAGGCGAAGGAATCTTAAGTTTGAAGGA							
Db	481	GGCTTAACCTGCTAGTTGATGAGTGGATAGGCGAAGGAATCTTAAGTTTGAAGGA							
Qy	1866	GGCTGAATGAAGAAATGTTTGAAGCTGATCGACAAAGTCAACTTGAACGGCAATTCAG							
Db	541	GGCTGAATGAAGAAATGTTTGAAGCTGATCGACAAAGTCAACTTGAACGGCAATTCAG							
Qy	1926	ATGATATATCATCTCAATGAACAGATCCGAATGTTGAACCTTACCGATATCATTTGCGA							
Db	601	ATGATATATCATCTCAATGAACAGATCCGAATGTTGAACCTTACCGATATCATTTGCGA							
Qy	1986	CACGAAGGTGCTTGTGACAGCTGCAATT	2015						
Db	661	CACGAAGGTGCTTGTGACAGCTGCAATT	690						
RESULT 7									
LOCUS	BQ405173								
DEFINITION	BQ405173	683 bp	mRNA	linear	EST 22-MAY-2002				
ACCESSION	GA_E00079E08f	Gossypium arboreum 7-10 dpa fiber library							
VERSION	arbozeum cDNA clone GA_E00079E08f,								
KEYWORDS	BQ405173								
SOURCE	BQ405173.1	GI:21092860							
ORGANISM	Gossypium arboreum								
REFERENCE	Gossypium arboreum								
AUTHORS	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.								
TITLE	1 (bases 1 to 683)								
JOURNAL	Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.								
COMMENT	An integrated analysis of the genetics, development, and evolution of the cotton fiber								
	Unpublished (2000)								
	Contact: Wing RA								
	Clemson University Genomics Institute								
	Clemson University								
	100 Jordan Hall, Clemson, SC 29634, USA								
	Tel: 864 656 7288								
	Fax: 864 656 4293								
	Email: rwing@clemson.edu								
	Total High Quality bases = 573								
	Seq primer: TAATACGACTCACTATAGG								
	High quality sequence start: 5								
	High quality sequence stop: 669.								
	Location/Qualifiers								
FEATURES									
SOURCE	1..683								
	/organism="Gossypium arboreum"								

```

/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0079S08f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV, Site_1: EcoRI, Site_2: XhoI"

```

Query Match	25.7%;	Score 674;	DB 13;	Length 683;
Best Local Similarity	99.1%;	Pred. No. 2,7e-116;		
Matches 677;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	180	GAAGCTCGCTAATGTGTGCATTTTGTGAAATATGAAAGGCTAGTCAGGAAGCGATCGTGT	239	
Db	1	GAAGCTCGCTAATGTGTGCATTTTGTGAAATATGAAAGGCTAGTCAGGAAGCGATCGTGT	60	
QY	240	GCCCTCCATGGGTTGCACTTGTCTGTTCGTCAGAGCCCTGGTGTGGAGATACATTAGAGT	299	
Db	61	GCCCTCCATGGGTTGCACTTGTCTGTTCGTCAGAGCCCTGGTGTGGAGATACATTAGAGT	120	
QY	300	GAATGTTCAAGCCCTTGTGTGAGGAACCTCAGTGTGATGATCCCACTCAAGGA	359	
Db	121	GAATGTTCAAGCCCTTGTGTGAGGAACCTCAGTGTGATGATCCCACTCAAGGA	180	
QY	360	AGAGCTGTGTGATGAAAGTTCAATGAAACTTTGTGTAATTTGAAATTTTGAAGCCCTT	419	
Db	181	AGAGCTGTGTGATGAAAGTTCAATGAAACTTTGTGTAATTTGAAATTTTGAAGCCCTT	240	
QY	420	CAACTCATCTTCCCCCGCCCACTCTTTCAAAATCCATTGTAATGATGTGAGTTCCT	479	
Db	241	CAACTCATCTTCCCCCGCCCACTCTTTCAAAATCCATTGTAATGATGTGAGTTCCT	300	
QY	480	AAATTCGCACCTTTGGGCAAAATTTGTCATGACAAAGAGACATGACACCTTTGCTCGA	539	
Db	301	AAATTCGCACCTTTGGGCAAAATTTGTCATGACAAAGAGACATGACACCTTTGCTCGA	360	
QY	540	ATTCTCTCAGAGTCCATTGTGCACAAAGGCAAGAACATGATGTTGAATGACAAATTCAGAA	599	
Db	361	ATTCTCTCAGAGTCCATTGTGCACAAAGGCAAGAACATGATGTTGAATGACAAATTCAGAA	420	
QY	600	CTTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGAGTATCTTGGTACCTTACCTCC	659	
Db	421	CTTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGAGTATCTTGGTACCTTACCTCC	480	
QY	660	TGAGACACCATGTGCGCAATTGCAACACCGGTCTCCAGAAATCGGTTTGGAAAGATGTG	719	
Db	481	TGAGACACCATGTGCGCAATTGCAACACCGGTCTCCAGAAATCGGTTTGGAAAGATGTG	540	
QY	720	GGGAGACACCGCAGAAAGCGGTGCTCGAGATGATCCAATCCTTTTGGATCTTTTGAAGGC	779	
Db	541	GGGAGACACCGCAGAAAGCGGTGCTCGAGATGATCCAATCCTTTTGGATCTTTTGAAGGC	600	
QY	780	AACGTATCTTTGACACCTTGAAGATTCCTTGGAGAAATCCCAATGATGATCAATGTTGT	839	
Db	601	AACGTATCTTTGACACCTTGAAGATTCCTTGGAGAAATCCCAATGATGATGATGATGAT	660	
QY	840	GATTTCTACTCTCCCAAGGATACT	862	
Db	661	GATTTCTACTCTCCCAAGGATACT	683	
RESULT 8				
LOCUS	CD486446	746 bp	mRNA	linear
LOCUS	CD486446			
DEFINITION	CHN2.2F08 Cotton Root and Hypocotyl Lambda ZIPLOX Library (CRH)			
	Gossypium hirsutum cDNA clone CHN2.2F08 5' similar to sucrose			
	synthase, mRNA sequence.			
ACCESSION	CD486446			
VERSION	CD486446.1	GI:31407411		

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
Gossypium hirsutum (upland cotton)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 746)
Dowd, C., Wilson, I. and McFadden, H.
Different Gene Expression Responses in Cotton Root and Hypocotyl
tissues during infection with Fusarium Wilt Disease
Unpublished (2003)
Contact: Caltriona Dowd, Helen McFadden
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry
Black Mountain Laboratories, Cnr Clunies Ross Street & Barry Drive,
Black Mountain, Canberra, ACT, 2601, Australia
Tel: 61 2 6246 4914, 6246 5377
Fax: 61 2 6246 5000
Email: Caltriona.Dowd@csiro.au, Helen.McFadden@csiro.au
Vector clipped sequences Bases 1-17 (GTGACCCACCGCTCCG): Sali
adaptec
Seq primer: M13 reverse primer
High quality sequence stop: 746.
Location/Qualifiers
1..746
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="DeltaEMERALD"
/db_xref="taxon:3635"
/clone="CRH2.2F08"
/tissue_type="root and hypocotyl tissues"
/dev_stage="5 day old seedlings"
/lab_host="Y1090(ZL)"
/clone_lib="Cotton Root and Hypocotyl Lambda ZIPLOX
Library (CRH)"
/note="Vector: Lambda ZIPLOX; Site 1: Sali; Site 2: NotI;
mRNA was prepared from root and hypocotyl tissues of the
cotton cultivar DeltaEMERALD. cDNA was synthesised from a
NotI-oligo(dT) primer/adaptor using the manufacturers
protocol (Life Technologies) and then ligated to a Sali
adaptor to facilitate directional cloning. The cDNA was
cloned into the Sali and NotI sites of the Lambda ZIPLOX
phage vector (Life Technologies). Constructed by Caltriona
Dowd and Helen McFadden."

ORIGIN

Query Match 25.3%; Score 664.8; DB 14; Length 746;
Best Local Similarity 97.2%; Pred. No. 1.4e-114;
Matches 697; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

1553 CCAAAATTCAACATTTTCCCTGCTGATATGAGATATTAATCTTCCCTTACACCGAAG 1612
Db 32 CCAAAATTCAACATTTTCCCTGCTGATATGAGATATTAATCTTCCCTTACACCGAAG 1612
1613 AGAAGCGGAGTTGAAGCATTTCCATCCGAGATGGAAGACCTTTTACACCAAGTTG 1672
Db 92 AGAAGCGGAGTTGAAGCATTTCCATCCGAGATGGAAGACCTTTTACACCAAGTTG 1672
1673 AGAATGAGAACACTATATGTGTCTCAATGACCGCAACGCAATTCGTTCACAAATGC 1732
Db 152 AGAATGAGAACACTATATGTGTCTCAATGACCGCAACGCAATTCGTTCACAAATGC 211
1733 CAAGGCTTGATGTCATCAAGACTTAAACCGGACTGTCAGTGTACCGCAAGACGCA 271
Db 212 CAAGGCTTGATGTCATCAAGACTTAAACCGGACTGTCAGTGTACCGCAAGACGCA 271
1793 AGTTCGCGAGTTGCTACCTCGTAGTTGATGATGAGATGGGAAAGAAATCAAG 1852
Db 272 AGTTCGCGAGTTGCTACCTCGTAGTTGATGATGAGATGGGAAAGAAATCAAG 331
1853 ATTGGAAGAGAGGCTGAATGAGAAATGTTTGTAGCTGACAGTCAACTGA 1912
Db 332 ATTGGAAGAGAGGCTGAATGAGAAATGTTTGTAGCTGACAGTCAACTGA 391

1913 ACGGCCAATTGAGATGATATCATCTCAAAATGAGACAGATCCGAATGTTGAACCTTACC 1972
Db 392 ACGGCCAATTGAGATGATATCATCTCAAAATGAGACAGATCCGAATGTTGAACCTTACC 451
1973 GATACATTTTGAGACAGGAAAGAGTCCCTTTGTACAGCTGCAATGATGAAAGCTTTGGAT 2032
Db 452 GATACATTTTGAGACAGGAAAGAGTCCCTTTGTACAGCTGCAATGATGAAAGCTTTGGAT 511
2033 TGACAGTTTGTGAGAGCAATGATCTTGGCGTTTGGCCAAATTCGCAACCTGTAAAGCTGAC 2092
Db 512 TGACAGTTTGTGAGAGCAATGATCTTGGCGTTTGGCCAAATTCGCAACCTGTAAAGCTGAC 571
2093 CAGCGAGATATTTGTCATCTGAGAAATCTGTTTCAACATTTGATCTTACCATGATC 2152
Db 572 CAGCGAGATATTTGTCATCTGAGAAATCTGTTTCAACATTTGATCTTACCATGATC 631
2153 AAGCTGCTGACATCTGCTGATTTCTTTGAAAAGTAAAGAAATCATCTGACGCG 2212
Db 632 AAGCTGCTGACATCTGCTGATTTCTTTGAAAAGTAAAGAAATCATCTGACGCG 690
2213 ATTAAGATCTCCCAAGAGGCTTGAAGGATAGAGAGAGATATACATGAAAGTTT 2269
Db 691 ATTAAGATCTCCCAAGAGGCTTGAAGGATAGAGAGAGATATACATGAAAGTTT 746

RESULT 9
BQ406766 682 bp mRNA linear EST 22-MAY-2002
LOCUS
DEFINITION
GA_E00098E07f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_E00098E07f, mRNA sequence.
BQ406766
VERSION
BQ406766.1 GI:21094453
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 682)
Wing, R.A., Friesch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A., and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rtwing@clemson.edu
Total High Quality bases = 571
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 676.
Location/Qualifiers
1..682
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_E00098E07f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 24.7%; Score 649.6; DB 13; Length 682;
Best Local Similarity 97.7%; Pred. No. 1e-111;
Matches 668; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

QY 804 GTTCTTGGGAGATCCCATGCTGTCAATGTTGTGATTTCTCACTCCCAACGATACTT 863
DB 1 GTTCTTGGGAGATCCCATGCTGTCAATGTTGTGATTTCTCACTCCCAACGATACTT 60
QY 864 CGCTCAAGACAAATGTTTGGGGTATCCGACACCGGTGGGAGGTTGTTTCACTTTGGA 923
DB 61 CGCTCAAGACAAATGTTTGGGGTATCCGACACCGGTGGGAGGTTGTTTCACTTTGGA 120
QY 924 TCAAGTCCGAGTTTGGAGAAATGAGATGCTCTCCGTATTAAGCAAGACATCAACAT 983
DB 121 TCAAGTCCGAGTTTGGAGAAATGAGATGCTCTCCGTATTAAGCAAGACATCAACAT 180
QY 984 CACCCCTCGAATCCTCATTTATTACTAGACTTTCTTCTGATGCTGTCGGAACAACATGCGG 1043
DB 181 CACCCCTCGAATCCTCATTTATTACTAGACTTTCTTCTGATGCTGTCGGAACAACATGCGG 240
QY 1044 TCAAGCACTTGAAGAAATGATCCGAAACAGACACTCGGATATTTCTTGAAGATCCCTTACG 1103
DB 241 TCAAGCACTTGAAGAAATGATCCGAAACAGACACTCGGATATTTCTTGAAGATCCCTTACG 300
QY 1104 AACAGAAAGGGAATTTGTCGAAAATGATCTCAAGATTTGAAAAGTCTGGCCATACTT 1163
DB 301 AACAGAAAGGGAATTTGTCGAAAATGATCTCAAGATTTGAAAAGTCTGGCCATACTT 357
QY 1164 GGAACCTTACACAGAGATGTTGCTCATGAAATCTTCAAAGATTGCAAGCGCCACA 1223
DB 358 GGAACCTTACACAGAGATGTTGCTCATGAAATCTTCAAAGATTGCAAGCGCCACA 417
QY 1224 TCTGATCATCGGAACCAAGGAGGCAATATGCTGCTCTTCTGTCGCAATAAAT 1283
DB 418 TCTGATCATCGGAACCAAGGAGGCAATATGCTGCTCTTCTGTCGCAATAAAT 477
QY 1284 AGGTGTCAACAGTGCACCATGCCCCATGCTTTGGAGAAACAAATATTCAGATTGCA 1343
DB 478 AGGTGTCAACAGTGCACCATGCCCCATGCTTTGGAGAAACAAATATTCAGATTGCA 537
QY 1344 TATCTATTGGAAGAAAGCTTTGAAGCAAAATACCATTTCTTGGCAATTTACAGTGAAT 1403
DB 538 TATCTATTGGAAGAAAGCTTTGAAGCAAAATACCATTTCTTGGCAATTTACAGTGAAT 597
QY 1404 TTTTGCATGAACCAATGATTTGATCATGACGATCTTTCCGGAATTTGACGAGAG 1463
DB 598 TTTTGCATGAACCAATGATTTGATCATGACGATCTTTCCGGAATTTGACGAGAG 657
QY 1464 CAAGGACACTGTTGCTCAATAGA 1487
DB 658 CANGACACTGTTGCTCAATAGA 681

RESULT 10
CF514627 951 bp mRNA linear EST 09-SEP-2003
LOCUS CF514627
DEFINITION Bud - CABUD Vitis vinifera cv. cabernet sauvignon (Clone 8)
Sequence.
ACCESSION CF514627
VERSION CF514627.1 GI:34546395
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 951)
AUTHORS Goes da Silva,F., Iandolo,A., Lim,H., Baek,J., Jones,K. and Cook,D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon',
berries at various developmental stages
JOURNAL Contact: Douglas Cook, PhD
COMMENT CABS Genome Facility
UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGGTACCGCATATGACC.
Location/Qualifiers
1. 951
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon (Clone 8)"
/db_xref="taxon:29760"
/clone="Cabud0005_1IF_H03"
/sex="hermaphrodite"
/dev_stage="Pre-bloom (10-11 days before bloom)"
/lab_host="DHSalpa"
/clone_1db="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"
/notes="Organ: Bud; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGGTATCAACGACGAGTGGCCATTAAGCGCGG-3' and
5'-ATCTAGAGCGGCGGCGCATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 24.1%; Score 633; DB 14; Length 951;
Best Local Similarity 79.7%; Pred. No. 1,2e-108;
Matches 760; Conservative 0; Mismatches 191; Indels 3; Gaps 1;
QY 719 GGGGTGACACCGGAGAACGGTGTCTCGAGATGATCCAACTCTTTGATGATCTTGTAGG 778
DB 1 GGGGTGATGATGCTGACGAGTGTCTGAGAGATGATTCATCTTTGACCTTTGAGG 60
QY 779 CAATGATGCTTGGACCTTGGAGAAAGTTCTTGGGAGAAATCCCATGGTGTTCATGTTG 838
DB 61 CTCCTGACCCCTGACACTTGAAGCAATTTCTTGGAGAAATCCCATGGTGTTCATGTTG 120
QY 839 TGATTTCACTCCCGACGATCTTCCCTCAAGCAATGTTTGGGTATCCGACACG 898
DB 121 TTATTTCTCCCCCATGCTACTTGTCTGACGAAATGTTCTGGGCTACCTGACACTG 180
QY 899 GTGGCAGGTTGTTTAACTTGAATCAAGTCGAGCTTTGGAGAAAGATGCTCTCC 958
DB 181 GTGGCAGGTTGTTTAACTTGAATCAAGTTCTGTCATGAGACATGAGATGCTTCTCC 240
QY 959 GTATTAAGCAACAGACTCAATCAACCCCTCGAATCTCATATTACTAGACTTCTTC 1018
DB 241 GTATTAAGCAACAGACTTGAATCACTCCCAAGATCATATTGTAAGTGAATCTTCC 300
QY 1019 CTGATGCTGTGGAACAACATGCGTCAACGACTTGAAGAAATGATCGGAACAGACAT 1078
DB 301 CTGATGCGGTAGGAGACATTTGCAACACGCGTATCGAGAAAGTTTATGGAACAGACAT 360
QY 1079 CGGATATTTCTGAGTACCTTCAAGCAAGAAAGGAATTTGTCGAAAATGATCTTCA 1138
DB 361 CAATCATCTTTCGAGTTCCCTTTGAACGTGAAGGGAATTTGTCACAAATGATCTTCA 420
QY 1139 GATTGAAAAGTGTGGCATACTTGGAAAACCTTACAGAGATGTTGCTCATGAAATCT 1198
DB 421 GATTG---AAGTGTGACATCTTGAAGAACTTACATGAGATGTTGCAAAAAGCTTG 477
QY 1199 CAAAAGTTGCAAGGACGACAGATGATCATCGAAAACNACGACGACGCAATATCG 1258
DB 478 CTACAGGCTACAGAACGACGACATTTTATCATTTGCAATTAACAGATGGAACAATCG 537
QY 1259 TCGCTCTTGTCTGACATTAATTAAGTGTACACAGTGCACCATGCGCCATGCTTTGG 1318

```

Db      538  TTGCTCTTGTGCTGCTATAGCTAGGGGTTACACAGTGCACCTAATCTCATGCTGCG 597
Qy      1319  AGAAGCAAAATATTCAGATTCAATCTATTTGGAAGAAGCTTGAACAATATTCAT 1378
Db      598  AGAAAAACAATATTCAGAAATCAATCTATTTGGAAGAATATTCAGAAATATTCAT 657
Qy      1379  TCTCTTGCAATTTTACAGTGTATCTTTTGGCAATAGCAATACAGATTTCATCACA 1438
Db      658  TCTATGCAAGTTTACAGTGTATCTTTTGGCAATAGCAATACAGATTTCATCACA 717
Qy      1439  GTACTTTCAGAAATTCAGAGCAAGACACTGTTGTCATATAGAGAGCCACTG 1498
Db      718  GTACTTTCAGAAATTCAGAGCAAGACACTGTTGTCATATAGAGAGCCACTG 777
Qy      1499  CTTTCACTCTTCTGCTGCTCTTCACTGCTGTTGTCATATAGAGAGCCACTG 1558
Db      778  GATTCAACATGCTGCTGCTGCTGTCATAGAGAGCCACTGTTTGAAGCCCAAT 837
Qy      1559  TCAACATGTTTCCCTGCTGCTGTCATAGAGAGCCACTGTTTGAAGCCCAAT 1618
Db      838  TCAACATGTTTCCCTGCTGCTGTCATAGAGAGCCACTGTTTGAAGCCCAAT 897
Qy      1619  GGAGGTTGAAGCATTTCCATCTGAGATCGAAGACCTTCTTACACCAAGTTG 1672
Db      898  TGAGGCTGAAGCCCTCCATCCAGAAATGGAAGACCTTCTTGAAGCCCTG 951

```

RESULT 11
BQ409552/c 656 bp mRNA linear EST 22-MAY-2002

DEFINITION GA_Ed0021E04r Gossypium arboreum 7-10 dpa fiber library Gossypium
LOCUS BQ409552
VERSION BQ409552.1 GI:21097239
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum

REFERENCE 1 (bases 1 to 656)
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)

JOURNAL
COMMENT Clemson University Genomics Institute
Contact: Wing RA
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

TITLE
JOURNAL
COMMENT Clemson University Genomics Institute
Contact: Wing RA
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

FEATURES
source
1..656
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0021E04r"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 24.0%; Score 630.2; DB 13; Length 656;
Best Local Similarity 98.6%; Pred. No. 4.4e-108;
Matches 646; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

```

Qy      1972  CGATACATTTTGGCAGACGAAGGTC -CTTTGTACAGCTGCATTTGATGAAGCTTTG 2030
Db      656  CGATACATTTTGGCAGACGAAGGTCCTTTGTACAGCTGCATTTGATGAAGCTTTG 597
Qy      2031  ATTGACAGTTTGGAGGCAATGACCTGCGGTTTSCCAATTCGCAACTGTAAAGGTTG 2090
Db      596  ATTGACAGTTTGGAGGCAATGACCTGCGGTTTSCCAATTCGCAACTGTAAAGGTTG 537
Qy      2091  ACCAGCCGAGATTATTCATAGGAAATCGTTTCAATTTGATTCATTCATGTTGA 2150
Db      536  ACCAGCCGAGATTATTCATAGGAAATCGTTTCAATTTGATTCATTCATGTTGA 477
Qy      2151  TCAAGCTGCTGACATCTCGTGATTTCTTTGAAAAGTGTAAAGATCATCTCACTG 2210
Db      476  TCAAGCTGCTGACATCTCGTGATTTCTTTGAAAAGTGTAAAGATCATCTCACTG 417
Qy      2211  GGATTAAGATCTCCAGAGAGGCTTGAACGAATGAGAGAGATATCATGGAAGATT 2270
Db      416  GGATTAAGATCTCCAGAGAGGCTTGAACGAATGAGAGAGATATCATGGAAGATT 357
Qy      2271  CTGAGAGACTATTTGACCTCTGACAGAGTGTATGATTCGAAAGCATGTTTCAACT 2330
Db      356  CTGAGAGACTATTTGACCTCTGACAGAGTGTATGATTCGAAAGCATGTTTCAACT 297
Qy      2331  TGAAGCCGCTGAGAGTGTGTTTACCTTGAGATGTTTATGCTTTAATGACCTTA 2390
Db      296  TGAAGCCGCTGAGAGTGTGTTTACCTTGAGATGTTTATGCTTTAATGACCTTA 237
Qy      2391  GCGTAATCACTTCATTTGAGAGAGATTAATGAACCTGTAAATGACATGAGGCGG 2450
Db      236  GCGTAATCACTTCATTTGAGAGAGATTAATGAACCTGTAAATGACATGAGGCGG 177
Qy      2451  TTTTCTTGAGAAATTAATCTGTTTGTAAATTTCAATTTGAGAGAGCTCTTTGTATT 2510
Db      176  TTTTCTTGAGAAATTAATCTGTTTGTAAATTTCAATTTGAGAGAGCTCTTTGTATT 117
Qy      2511  CATCTGTCTTTCTTTCTTTCTTTTTCGCGGCAATGTTTGAACATGAGGTTGCGCC 2570
Db      116  CATCTGTCTTTCTTTCTTTCTTTTTCGCGGCAATGTTTGAACATGAGGTTGCGCC 57
Qy      2571  CGTCAATTCAGTTAATATGATGATTTTGTTCATTTTCAAAAAA 2625
Db      56  CGTCAATTCAGTTAATATGATGATTTTGTTCATTTTCAAAAAA 2

```

RESULT 12
BQ415686/c 637 bp mRNA linear EST 22-MAY-2002

DEFINITION GA_Ed0102A07r Gossypium arboreum 7-10 dpa fiber library Gossypium
LOCUS BQ415686
VERSION BQ415686
KEYWORDS EST.
SOURCE Gossypium arboreum

REFERENCE 1 (bases 1 to 637)
Wing, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)

JOURNAL
COMMENT Clemson University Genomics Institute
Contact: Wing RA
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288

TITLE
JOURNAL
COMMENT Clemson University Genomics Institute
Contact: Wing RA
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288

Fax: 864 656 4293
Email: rwing@clmson.edu
Total High Quality bases = 539
Seq primer: TAATGACTCATATAGCG
High quality sequence start: 3
High quality sequence stop: 632.
Location/Qualifiers

FEATURES

Source

1. 637
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0102A07r"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.7%; Score 623.2; DB 13; Length 637;

Best Local Similarity 98.7%; Pred. No. 9.2e-107;
Matches 628; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1989 GAAAGTGCCTTTGACAGCTGACCTGATGAGACCTTTGGATTGACAGTTGTGAGGC 2048
636 GAAAGTGCCTTTGACAGCTGACCTGATGAGACCTTTGGATTGACAGTTGTGAGGC 577
2049 AATGACTGCGGTTGCCAACAATTGCCAACCTGTACCGGTGACCGACCGAGATTATGT 2108
576 AATGACTGCGGTTGCCAACAATTGCCAACCTGTACCGGTGACCGAGATTATGT 517
2109 CCATGGAAATCTGGTTCAACATTGATCTTACCGATGATCAAGCTGCTGACATCT 2168
516 CCATGGAAATCTGGTTCAACATTGATCTTACCGATGATCAAGCTGCTGACATCT 457
2169 CGTCGATTTCTTTGAAAAGTGAAGAAAGATCCATCTCACTGGGTAAAGATCTCCAAAG 2228
456 CGTCGATTTCTTTGAAAAGTGAAGAAAGATCCATCTCACTGGGTAAAGATCTCCAAAG 397
2229 AGGCTTGAAGCAAGAGAGAGAGATATATCATGAGAGATTACTCGGAGAGATTTGAC 2288
396 AGGCTTGAAGCAAGAGAGAGAGATATATCATGAGAGATTACTCGGAGAGATTTGAC 337
2289 CCTGACAGAGTGAATGATCTGAGACATGTTTCCAACTTGAACCCCGTAGAGTGC 2348
336 CCTGACAGAGTGAATGATCTGAGACATGTTTCCAACTTGAACCCCGTAGAGTGC 277
2349 TCGTTACCTTGAGATGTTTATGCTCTTAAAGTACCGTAAGCTGCTGAATCAATTCA 2408
276 TCGTTACCTTGAGATGTTTATGCTCTTAAAGTACCGTAAGCTGCTGAATCAATTCA 217
2409 GCGAGAGAGTGAATGAACTGTAAATTAACATGGGCGCGTTTCTTGGAGAAATAT 2468
216 GCGAGAGAGTGAATGAACTGTAAATTAACATGGGCGCGTTTCTTGGAGAAATAT 157
2469 ATTCTGTTTGAATTTCAATTTGAGAAAGCTCTTGTATTTCACTTGTCTTTCTTT 2528
156 ATTCTGTTTGAATTTCAATTTGAGAAAGCTCTTGTATTTCACTTGTCTTTCTTT 97
2529 TCGTTTTCGCGGCAATTTGTTGAACATGGGCTGTGCGCGCGTCAATTCAGTTAAAT 2588
96 TCGTTTTCGCGGCAATTTGTTGAACATGGGCTGTGCGCGCGTCAATTCAGTTAAAT 37
2589 ATGTGACTTTTGTCTTTTCAAAAAAATTTTTTTTTT 2624
36 ATGTGACTTTTGTCTTTTCAAAAAAATTTTTTTTTT 1

RESULT 13

AI731292 AT731292 664 bp mRNA linear EST 11-JUN-1999

DEFINITION BN1GH19102 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U73588) sucrose synthase (Gossypium hirsutum), mRNA sequence.
ACCESSION AI731292
VERSION AI731292.1 GI:5050144
KEYWORDS EST.
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 664)
Blewitt M., Marx E.C., Davy D.F. and Burr B.
ESTs from developing cotton fiber
Unpublished (1999)
JOURNAL
TITLE
AUTHORS
COMMENT

Brookhaven National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burr@bnl.bnl.gov
Seq primer: T3 Primer.
Location/Qualifiers

FEATURES

Source

1. 664
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/tissue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="X11-Blue"
/clone_lib="Six-day Cotton fiber"
/note="Vector: pBluescript II KS+"

ORIGIN

Query Match 23.7%; Score 623.2; DB 9; Length 664;

Best Local Similarity 97.7%; Pred. No. 9.1e-107;
Matches 642; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

1713 GCCAATTCGTGTCACAATGCCAAGCTTGATCGTGCAGAACCTTAACCGGACTCGTGA 1772
1 GCCAATTCGTGTCACAATGCCAAGCTTGATCGTGCAGAACCTTAACCGGACTCGTGA 60
1773 GTGTCGCGCAAGAACCCAAAGTTGCGTGAAGTGGCTAACCTCGTGAAGTGGTGA 1832
61 GTGTCGCGCAAGAACCCAAAGTTGCGTGAAGTGGCTAACCTCGTGAAGTGGTGA 120
1833 TAGCGGAAGGAATCTAAAGATTGGAGAGAGGCTGAATGAAGAAATGTTGAGCT 1892
121 TAGCGGAAGGAATCTAAAGATTGGAGAGAGGCTGAATGAAGAAATGTTGAGCT 180
1893 GATGACAAGTACAACCTTGAACCGCCCAATTCAGATGATATCATCTCAAAATGAACAGAT 1952
181 GATGACAAGTACAACCTTGAACCGCCCAATTCAGATGATATCATCTCAAAATGAACAGAT 240
1953 CCGAAATGTTGAACCTTACCGATATCATTTGGAGACAGAAAGGTCCTTTGTACAGCTGC 2012
241 CCGAAATGTTGAACCTTACCGATATCATTTGGAGACAGAAAGGTCCTTTGTACAGCTGC 300
2013 ATTGTAAAGACCTTGGATTGACAGTTGGAGGCAATGACTGGGTTGGCAACTT 2072
301 ATTGTAAAGACCTTGGATTGACAGTTGGAGGCAATGACTGGGTTGGCAACTT 360
2073 CGCAACCTGTAAACGGTGAACCGCCAGATTATTTGTCATGAGGAAATCTGTTCAACAT 2132
361 CGCAACCTGTAAACGGTGAACCGCCAGATTATTTGTCATGAGGAAATCTGTTCAACAT 420
2133 TGATCTTACCATGATGATCAAGCTGTGACATATCTGTGATTTCTTTGAAAAGTGA 2192
421 TGATCTTACCATGATGATCAAGCTGTGACATATCTGTGATTTCTTTGAAAAGTGA 480
2193 GAAAGATCCATCTCACTGGGATGAAGATCTCCAAAGAGGCTTGAAGGAATAGAGGAA 2252

Db 481 GAAAGATTCATTCACCTGGGATTAAGTATCCCAAGAGGCTTGAACGATTAAGAGAA 540

Qy 2253 GTATACATGAGAGATTACTCGAGAGACTATTGACCCCTGACAGAGGTGTATGATTCG 2312

Db 541 ATATACATGAGAGATTACTCGAGAGACTATTGACCCCTGACAGAGGTGTATGATTCG 600

Qy 2313 GAGAGATGTTT-CCAACTTGAAAGCCGCTGAGAGTGTGCTTACCTTGAGATGTTT 2368

Db 601 GAAACATGTTTCCCACTTGAAAGCCGCTGAGAGTGTGCTTACCTTGAAATGTTT 657

RESULT 14

BG442235 642 bp mRNA linear EST 15-MAR-2001

LOCUS GA_Ea001E08f Gossypium arboreum 7-10 dpa fiber library Gossypium

DEFINITION BG442235

ACCESSION BG442235

VERSION BG442235.1 GI:13351887

KEYWORDS EST.

SOURCE Gossypium arboreum

ORGANISM Gossypium arboreum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 642)

AUTHORS Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT Contact: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7286

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: TAATACGACTCACAATAGG

High quality sequence stop: 640.

FEATURES

Source

1. 642

/organism="Gossypium arboreum"

/mol_type="mRNA"

/strain="AKA"

/cultivar="8400"

/db_xref="taxon:29729"

/clone="GA_Ea001E08f"

/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

/lab_host="E. coli"

/clone_lhb="Gossypium arboreum 7-10 dpa fiber library"

/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.7%; Score 623; DB 12; Length 642;

Best Local Similarity 99.1%; Pred. No. 1e-106;

Matches 637; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1916 GCCAATTCAGATGATATCATCTCAATGAAGAGATCGAATGTTGAATCCGAT 1975

Db 1 GCCAATTCAGATGATATCATCTCAATGAAGAGATCGAATGTTGAATCCGAT 59

Qy 1976 ACATTTGGAGACGAAAGGTGCTTTGTACAGCCCTGCTTGTATGAAGCCCTTTGATTTGA 2035

Db 60 ACATTTGGAGACGAAAGGTGCTTTGTACAGCCCTGCTTGTATGAAGCCCTTTGATTTGA 119

Qy 2036 CAGTTTGAGAGCAATGACTTGGGTTTCCAAACATTCGCAACCTGTAAACGTTGACAG 2095

Db 120 CAGTTTGAGAGCAATGACTTGGGTTTCCAAACATTCGCAACCTGTAAACGTTGACAG 179

Qy 2096 CCGAGATTAATGCTCCAGGAAATCTGGTTTCAACATTTGATCTTACCATGTTATCAAG 2155

Db 180 CTGAGATTAATGCTCCAGGAAATCTGGTTTCAACATTTGATCTTACCATGTTATCAAG 239

Qy 2156 CTGCTGACATATCTGTCGATTTCTTTGAAAAGTGAAGAGATCATCTCACTGGATA 2215

Db 240 CTGCTGACATATCTGTCGATTTCTTTGAAAAGTGAAGAGATCATCTCACTGGATA 299

Qy 2216 AGATCTCCCAAGAGGCTTGAACGAATGAGAGAACTATACATGAAATTTACTCGG 2275

Db 300 AGATCTCCCAAGAGGCTTGAACGAATGAGAGAACTATACATGAAATTTACTCGG 359

Qy 2276 AGAGACTATTAACCTGACAGAGTGTATGATTTCTGGAACATGTTTCCAACCTGAAAC 2335

Db 360 AGAGACTATTAACCTGACAGAGTGTATGATTTCTGGAACATGTTTCCAACCTGAAAC 419

Qy 2336 GCCGTGAGAGTGTGCTGATACCTTGAGATGTTTATGCTTAAAGACCGTAAGCTGCTG 2395

Db 420 GCCGTGAGAGTGTGCTGATACCTTGAGATGTTTATGCTTAAAGACCGTAAGCTGCTG 479

Qy 2396 AATCAGTTCCATTGGCAGAGAGATTAATGAACCTGTTAAATTAACATTTGGCCGCTTTT 2455

Db 480 AATCAGTTCCATTGGCAGAGAGATTAATGAACCTGTTAAATTAACATTTGGCCGCTTTT 539

Qy 2456 CTGGAGAAATTAATTTCTGTTTGTATTAATTCAATTGAGAACTCTTGTATTTATCT 2515

Db 540 CTGGAGAAATTAATTTCTGTTTGTATTAATTCAATTGAGAACTCTTGTATTTATCT 599

Qy 2516 TGTCTTTTCCCTTCTCTTTTTCGCCGCGCATTTGTTGAACATG 2558

Db 600 TGTCTTTTCCCTTCTCTTTTTCGCCGCGCATTTGTTGAACATG 642

RESULT 15

BE054845 871 bp mRNA linear EST 07-MAR-2001

LOCUS BE054845

DEFINITION GA_Ea0032D04f Gossypium arboreum 7-10 dpa fiber library Gossypium

ACCESSION BE054845

VERSION BE054845.2 GI:13245895

KEYWORDS EST.

SOURCE Gossypium arboreum

ORGANISM Gossypium arboreum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 871)

AUTHORS Wing,R.A., Fritsch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT On Jun 8, 2000 this sequence version replaced gi:8381998.

Contact: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7286

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: TAATACGACTCACAATAGG

High quality sequence stop: 581.

FEATURES

Source

1. 871

/organism="Gossypium arboreum"

/mol_type="mRNA"

/strain="AKA"

/cultivar="8400"

/db_xref="taxon:29729"

/clone="GA_Ea0032D04f"

/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

/lab_host="E. coli"

/clone_lhb="Gossypium arboreum 7-10 dpa fiber library"

/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.6%; Score 620.2; DB 10; Length 871;
Best Local Similarity 86.0%; Pred. No. 3e-106;
Matches 734; Conservative 0; Mismatches 113; Indels 6; Gaps 4;

```
OY 1460 GAAGGAAGACACCTTGTGTGTAATCGAGGCCACACCTGCTTTCATCTTCCTGCTCT 1519
    |||
Db 1 GAAGCAAGACACCTGTGTGTGTAATCGAGGCCACACCTGCTTTCATCTTCCTGCTCT 60
    |||

OY 1520 ACCGTGTGTATCATGTATCATGTGTGTGTATCCCAATTCAACATGTTCCTCGGTG 1579
    |||
Db 61 ACCGTGTGTATCATGTATCATGTGTGTGTATCCCAATTCAACATGTTCCTCGGTG 120
    |||

OY 1580 CTGATATGAGATATATCTTCCTTACACCGAAGAGAGCGAGGTTGAAGCATTTCCATC 1639
    |||
Db 121 CTGATATGAGATATATCTTCCTTACACCGAAGAGAGCGAGGTTGAAGCATTTCCATC 180
    |||

OY 1640 CTGAGATGGAAGACCTTCTTTACACCAAGTTGAAGATGAAGAACATTATGTGTCTCA 1699
    |||
Db 181 CTGAGATGGAAGACCTTCTTTACACCAAGTTGAAGATGAAGAACATTATGTGTCTCA 240
    |||

OY 1700 ATGACCGCAACAAAGCCAAATTCGTACAAATGCAAGGCTTGATCGTCAAGAACTTAA 1759
    |||
Db 241 ATGACCGCAACAAAGCCAAATTCGTACAAATGCAAGGCTTGATCGTCAAGAACTTAA 300
    |||

OY 1760 CCGGACTCGTGTGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATCTGCTAG 1819
    |||
Db 301 CCGGACTCGTGTGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATCTGCTAG 360
    |||

OY 1820 TTGTAGTGTGTGATATGAGGAAAGAAATCTAAAGATTTGGAAGAGAGGCTGAATGAAGA 1879
    |||
Db 361 TTGTAGTGTGTGATATGAGGAAAGAAATCTAAAGATTTGGAAGAGAGGCTGAATGAAGA 420
    |||

OY 1880 AAATGTTGAGCTGATCGACAAAGTACAACTTGAAGGCAATTCAGATGATATCATCTC 1939
    |||
Db 421 AAATGTTGAGCTGATCGACAAAGTACAACTTGAAGGCAATTCAGATGATATCATCTC 480
    |||

OY 1940 AAATGAACAGAAATCCGAATGTTGAACCTTTACCGATACATTTGCGACACGAAAGGTGCT 1999
    |||
Db 481 AAATGAACAGAAATCCGAATGTTGAACCTTTACCGATACATTTGCGACACGAAAGGTGCT 539
    |||

OY 2000 TTGTACAGCTGATGATATGAGGCTTTGATTTGACAGTTGTGAGGCAATGACTTGGC 2059
    |||
Db 540 TTGTACAGCTGATGATATGAGGCTTTGATTTGACAGTTGTGAGGCAATGACTTGGC 599
    |||

OY 2060 GTTTCACAACATTC-GCAACCTGTACGCTGGAACGACGAGATTATGTCATGGGAAA 2118
    |||
Db 600 GATTGGCCACATTCACGCAATGCTGACCGAGGACCAAGCTGGAATATTGACCATGGGAAA 659
    |||

OY 2119 TCTGTGTTCAACATGATCTTTACCATGCTGATCAAGCTGTCATACCTGCTGATTTTC 2178
    |||
Db 660 TGTGGTTAAACATGGGTTCTTTACATGGGATGAGGAAAGTAACTTCCGAAAGATG 719
    |||

OY 2179 TTTGAAAAGTGTAAAGAAATCATCTCATCTGGGATA--AGATCTCCCAAGAGGCTTG 2235
    |||
Db 720 TTTTAAAGAGGTGTGAAAAATTCATGAGCTTGAATATATATCAACAAGAGGCTTG 779
    |||

OY 2236 AAACGAATAGAGAGATATCATGAAAGATTACT-CGGAGAGACTATTGACCTGAC 2294
    |||
Db 780 AGCGAGTAAATATTAATAATTCATAGCAGAGGTTCTAGGAAAACACATAGAACCCGAC 839
    |||

OY 2295 AGGAGTGTATGA 2307
    |||
Db 840 AGGAGAGAGGGA 852
    |||
```

Search completed: June 5, 2004, 08:30:34
Job time : 9938 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:46:39 ; Search time 61 Seconds
(without alignments)
3733.333 Million cell updates/sec

Title: US-10-003-405-2

Sequence: 1 MAEALTRVSLRERLDEL.....EMFALRYKRLAEVPLAE 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	AAU97898	AAU97898 Cotton su
2	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
3	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
4	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
5	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
6	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
7	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
8	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
9	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
10	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
11	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
12	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
13	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
14	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
15	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
16	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
17	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
18	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
19	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
20	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
21	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
22	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
23	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
24	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
25	3643.5	85.8	805	AAU97898	AAU97898 Cotton su

26	2961	69.7	802	AAE28499	AAE28499 Corn suc
27	2875	67.7	797	ABR39586	ABR39586 A. thalia
28	2866.5	67.5	805	ABR39582	ABR39582 A. thalia
29	2866.5	67.5	805	ABR39585	ABR39585 A. thalia
30	2841	66.9	766	AAE66222	AAE66222 Sucrose-s
31	2312.5	54.5	942	ABR91573	ABR91573 Herbicida
32	2184.5	51.4	786	ADC07862	ADC07862 Rice prot
33	2147.5	50.6	798	ADC07860	ADC07860 Rice prot
34	2133	50.2	843	ABR93633	ABR93633 Herbicida
35	1747	41.1	806	AAW53103	AAW53103 Anabaena
36	1440	33.9	514	ADC07854	ADC07854 Rice prot
37	1383	32.6	395	ABG69054	ABG69054 Amino aci
38	1174.5	27.7	348	AAU16313	AAU16313 Pinus rad
39	879	20.7	225	ABG69052	ABG69052 Amino aci
40	707	16.6	198	AAU85667	AAU85667 Plant col
41	698	16.4	204	ABW73686	ABW73686 DNA clone
42	643.5	15.2	217	AAU16309	AAU16309 Eucalyptu
43	621	14.6	217	AAU16312	AAU16312 Pinus rad
44	603	14.2	242	ABG69053	ABG69053 Amino aci
45	574.5	13.5	158	AAU16284	AAU16284 Pinus rad

ALIGNMENTS

AAU97898	AAU97898 standard; protein; 806 AA.
AC	AAU97898;
DT	13-AUG-2002 (first entry)
DE	Cotton sucrose synthase Susy protein.
XX	Cotton; Sucrose synthase; Susy; fibre; seed; transgenic; plant; enzyme.
XX	Gossypium hirsutum.
XX	Key Location/Qualifiers
XX	Misc-difference 414
XX	/label= Unknown
XX	/note= "Encoded by NAC"
XX	MO200245485-A1.
XX	13-JUN-2002.
XX	07-DEC-2001; 2001MO-AU001580.
XX	08-DEC-2000; 2000AU-00001975.
XX	08-DEC-2000; 2000US-0251852P.
XX	(CSIR) COMMONWEALTH SCI & IND RES ORG.
XX	Llewellyn D, Furbank R, Ruan Y;
XX	WPI; 2002-463779/49.
XX	N-PSDB; ABX52880.
XX	Altering fibre development or properties of a fibre producing plant by
XX	modulating sucrose synthase activity and/or expression in such plants,
XX	useful for enhancing fibre yield and quality and for increasing seed
XX	size.
XX	Claim 5; Page 57-60; 62pp; English.
XX	This invention relates to a novel method for altering fibre development
XX	or properties of a fibre producing plant by modulating sucrose synthase
XX	(Susy) activity and/or expression in such plants. The invention also
XX	comprises a fibre producing plant comprising a chimaeric gene in its
XX	genome, the seeds of the plant and fibre isolated from the plant. The
XX	method is useful for altering fibre development or properties of a fibre

CC producing plant like cotton plant. Therefore, the method is useful for
 CC enhancing fibre yield, enhancing fibre quality and for increasing seed
 CC size in a fibre producing plant. The present sequence represents the
 CC cotton sucrose synthase (Susy) protein used to create the transgenic
 CC plant of the invention

XX Sequence 806 AA:

Query Match 100.0%; Score 4245; DB 5; Length 806;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAERALTFRVSHLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
Db 1 MAERALTFRVSHLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
Qy 61 KLANGAFPEVLKASOEAVLPWPVALAVRPPGWERYRNVNVALVVEELVAAYLHFKE 120
Db 61 KLANGAFPEVLKASOEAVLPWPVALAVRPPGWERYRNVNVALVVEELVAAYLHFKE 120
Qy 121 ELVDGSSNGNFVLELDEPFNSPPRPTLSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDEPFNSPPRPTLSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Qy 181 FLRVHCHKGKMMNDRIQNLNALQHVLRKAEEYLGTLPPETPCAFEPHRRFOEIGLERGW 240
Db 181 FLRVHCHKGKMMNDRIQNLNALQHVLRKAEEYLGTLPPETPCAFEPHRRFOEIGLERGW 240
Qy 241 GDTAERVLEMIQLLLDLLEATDPCTLEKFLGRIMVFNVVVLTTHGYFAQONVLYGYPDTG 300
Db 241 GDTAERVLEMIQLLLDLLEATDPCTLEKFLGRIMVFNVVVLTTHGYFAQONVLYGYPDTG 300
Qy 301 GQVVVYILDQVRLAENEMLRIRKQOGLNTTPRILITITRLIPAVGTCGRLKXYGTEHS 360
Db 301 GQVVVYILDQVRLAENEMLRIRKQOGLNTTPRILITITRLIPAVGTCGRLKXYGTEHS 360
Qy 361 DILRVFPTEKGIYRKWISREKWPYLETYTEDVAHESIKELHGTDDLIGNSDGNIV 420
Db 361 DILRVFPTEKGIYRKWISREKWPYLETYTEDVAHESIKELHGTDDLIGNSDGNIV 420
Qy 421 ASLAAHKLGVTCITAAHLEKTKYPPDSIDIVYKKLEDEKHSCEOPTADLFAMNHDTFIITS 480
Db 421 ASLAAHKLGVTCITAAHLEKTKYPPDSIDIVYKKLEDEKHSCEOPTADLFAMNHDTFIITS 480
Qy 481 TFOEIASGKOTVGOYESHFAFLPGLYVYVGVGIDVDFKFNIVSPGADMEIYFYTEER 540
Db 481 TFOEIASGKOTVGOYESHFAFLPGLYVYVGVGIDVDFKFNIVSPGADMEIYFYTEER 540
Qy 541 RLKGFHPEIEDLLYTKVNEBHLCVLDRNRPILFTMPRLDRVXNLTGLVEMCGNPKLR 600
Db 541 RLKGFHPEIEDLLYTKVNEBHLCVLDRNRPILFTMPRLDRVXNLTGLVEMCGNPKLR 600
Qy 601 ELAULVVGSGRRKESKLEKAEKMKFELIDKYNLNGORWISSQNRIRANVELRYXI 660
Db 601 ELAULVVGSGRRKESKLEKAEKMKFELIDKYNLNGORWISSQNRIRANVELRYXI 660
Qy 661 CDTGAFVOPALYEAFFGLTVVEAMTCGLPTATCGGPAEITIVGKSGSFNIDPHGDOAA 720
Db 661 CDTGAFVOPALYEAFFGLTVVEAMTCGLPTATCGGPAEITIVGKSGSFNIDPHGDOAA 720
Qy 721 DILVDFPEKCKDPSHMDKISQGLRIEKKYTKWYISERLLTLTGYYGFWKHSVNLERR 780
Db 721 DILVDFPEKCKDPSHMDKISQGLRIEKKYTKWYISERLLTLTGYYGFWKHSVNLERR 780
Qy 781 ESRRYLEMFYALKTRKLAESVPLAE 806
Db 781 ESRRYLEMFYALKTRKLAESVPLAE 806

```

RESULT 2
 AAB16282
 ID AAB16282 standard; protein; 805 AA.
 XX

```

AC AAB16282;
XX 31-OCT-2000 (first entry)
DT 31-OCT-2000 (first entry)
XX
DE Eucalyptus grandis sucrose synthase protein sequence SEQ ID NO:44.
XX
KM Eucalyptus grandis; pinus radiata; Monterey pine; modification;
XX plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant.
XX
OS Eucalyptus grandis.
XX
PN MO200022092-A2.
XX
PD 20-Apr-2000.
XX
PF 08-OCT-1999; 99NO-NZ000169.
XX
PR 13-OCT-1998; 98US-00170862.
XX
PR 11-AUG-1999; 99US-0148426P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR MPI; 2000-339328/29.
XX
DR N-PSDB; AAA67087.
XX
XX
XX New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant.
XX
XX Claim 17; Page 58-60; 301pp; English.
XX
XX
XX The present invention describes isolated polynucleotides (PN) comprising
XX a sequence selected from one of 835 nucleotide sequences given in
XX AAA67073 to AAA67907, their (reverse) complements, sequences producing an
XX expectation (E) value of 0.01 or less compared to the 835 sequences,
XX sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
XX 835 sequences or sequences that are degenerately equivalent or allelic to
XX the 835 sequences. The polynucleotides are used to modify the activity of
XX a polypeptide involved in a polysaccharide biosynthetic pathway in the
XX plant. They are especially used to modulate or alter the polysaccharide
XX content, composition or structure of the plant. AAB16268 to AAB16340 are
XX proteins encoded by some of the polynucleotide sequence given in the
XX present invention
XX
XX Sequence 805 AA:
XX
XX Query Match 85.8%; Score 3643.5; DB 3; Length 805;
XX Best Local Similarity 85.0%; Pred. No. 7.7e-316;
XX Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;
Qy 1 MAERALTFRVSHLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
Db 1 MAERALTFRVSHLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
Qy 61 KLANGAFPEVLKASOEAVLPWPVALAVRPPGWERYRNVNVALVVEELVAAYLHFKE 120
Db 61 KLANGAFPEVLKASOEAVLPWPVALAVRPPGWERYRNVNVALVVEELVAAYLHFKE 120
Qy 121 ELVDGSSNGNFVLELDEPFNSPPRPTLSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDEPFNSPPRPTLSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Qy 181 FLRVHCHKGKMMNDRIQNLNALQHVLRKAEEYLGTLPPETPCAFEPHRRFOEIGLERGW 240
Db 181 FLRVHCHKGKMMNDRIQNLNALQHVLRKAEEYLGTLPPETPCAFEPHRRFOEIGLERGW 240
Qy 241 GDTAERVLEMIQLLLDLLEATDPCTLEKFLGRIMVFNVVVLTTHGYFAQONVLYGYPDTG 300
Db 241 GDTAERVLEMIQLLLDLLEATDPCTLEKFLGRIMVFNVVVLTTHGYFAQONVLYGYPDTG 300

```

QY 301 GGVVYILDQVRLAENEMLRIRIKOQAGNITPRILITRLLPDAVGTTCGOREKXYGTHS 360
 DB 301 GGVVYILDQVRLAENEMLRIRIKOQAGNITPRILITRLLPDAVGTTCGOREKXYGTHS 360
 QY 361 DILRVFPTEKGIYRWKISREPEKWPYLETYTEDEVAAHESKELAGTDPDLIIGNSXGDNIV 420
 DB 361 DILRVFPTEKGIYRWKISREPEKWPYLETYTEDEVAAHESKELAGTDPDLIIGNSXGDNIV 420
 QY 421 ASLAAKLGVTQCTIAHALEKTKYPSDSIYWKULEDKTHFSQCFADLIPANNHTDPIITTS 480
 DB 421 ASLAAKLGVTQCTIAHALEKTKYPSDSIYWKULEDKTHFSQCFADLIPANNHTDPIITTS 480
 QY 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
 DB 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
 QY 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
 DB 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
 QY 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKYNLNGQFRWISSQNNRIRNVELYRYI 660
 DB 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKYNLNGQFRWISSQNNRIRNVELYRYI 660
 QY 661 CDTKGAFFQPALYEAFGITVVEAMTCGLPTPATCNGGPAEIIYVHGSGSPNIDPYHGDAA 720
 DB 661 CDTKGAFFQPALYEAFGITVVEAMTCGLPTPATCNGGPAEIIYVHGSGSPNIDPYHGDAA 720
 QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTTLGVYGFWMKXSNLEBR 780
 DB 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTTLGVYGFWMKXSNLEBR 780
 QY 781 ESRRLYEMFVALKTRKLAESVPLAE 806
 DB 781 ESRRLYEMFVALKTRKLAESVPLAE 806
 QY 780 ESRRLYEMFVALKTRKLAESVPLAE 805
 DB 780 ESRRLYEMFVALKTRKLAESVPLAE 805

RESULT 3

AAB16336 standard; protein; 805 AA.

AAB16336;

31-OCT-2000 (first entry)

Eucalyptus grandis sucrose synthase protein sequence SEQ ID NO:144.

Eucalyptus grandis; pinus radiata; Monterey pine; modification;

plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;

transgenic plant.

Eucalyptus grandis.

MO300022092-A2.

20-APR-2000.

08-OCT-1999; 99WO-NZ000169.

13-OCT-1998; 98US-00170862.

PR 11-AUG-1999; 99US-0148426P.

(GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Bloksberg LN;

WPI: 2000-339328/29.

N-PSDB; AAA67143.

New genes encoding proteins involved in a plant polysaccharide biosynthetic pathway, useful for modulating or altering the

PT polysaccharide content, composition or structure of the plant.

XX Claim 17; Page 108-110; 301pp; English.

CC The present invention describes isolated polynucleotides (PN) comprising
 CC a sequence selected from one of 835 nucleotide sequences given in
 CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an
 CC expectation (E) value of 0.01 or less compared to the 835 sequences.
 CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
 CC 835 sequences or sequences that are degenerately equivalent or allelic to
 CC the 835 sequences. The polynucleotides are used to modify the activity of
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
 CC plant. They are especially used to modulate or alter the polysaccharide
 CC content, composition or structure of the plant. AAB1636 to AAB16340 are
 CC proteins encoded by some of the polynucleotide sequence given in the
 CC present invention
 CC
 SQ Sequence 805 AA;

Query Match 85.8%; Score 3643.5; DB 3; Length 805;
 Best Local Similarity 85.0%; Pred. No. 7.7e-316;
 Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

QY 1 MAERALTREHSLRELDLTLAHRNEIILALSREKGGKGLIOHQIILEFEALPEENRK 60
 DB 1 MADRLTSHSLRERLDLTLAHRNDIYAFISRVAKKGLIQHQIIFAEFAISESRA 60
 QY 61 KLANGAFPEVILKASOEALVLPWVALAVRPPGWERYIRVNVHVLVEELVAYELHFE 120
 DB 61 KLNGAFGEVILKSTGEALVSPWVALAVRPPGWERYIRVNVHVLVEELVAYELHFE 120
 QY 121 ELVNGSSNGNVLDELDEFPNSSFPRPTLSSIGNGVEFLNRHLSAKLPHKSHAPLE 180
 DB 121 ELVNGSSNGNVLDELDEFPNSSFPRPTLSSIGNGVEFLNRHLSAKLPHKSHAPLE 180
 QY 181 FLRVHCHGKMMMLNDRIQNLNALOHVLRKAEVYIGTLPETPCADEFERFOEIGLERGM 240
 DB 181 FLVHCHGKMMMLNDRIQNLNALOHVLRKAEVYIGTLPETPCADEFERFOEIGLERGM 240
 QY 241 GDTAERVLEMIQLLDLLEADPTCTLEKFLGRIPWNVNVLTPHGYFAQDNVLGYPTG 300
 DB 241 GDTAERVLEMIQLLDLLEADPTCTLEKFLGRIPWNVNVLTPHGYFAQDNVLGYPTG 300
 QY 301 GGVVYILDQVRLAENEMLRIRIKOQAGNITPRILITRLLPDAVGTTCGOREKXYGTHS 360
 DB 301 GGVVYILDQVRLAENEMLRIRIKOQAGNITPRILITRLLPDAVGTTCGOREKXYGTHS 360
 QY 361 DILRVFPTEKGIYRWKISREPEKWPYLETYTEDEVAAHESKELAGTDPDLIIGNSXGDNIV 420
 DB 361 DILRVFPTEKGIYRWKISREPEKWPYLETYTEDEVAAHESKELAGTDPDLIIGNSXGDNIV 420
 QY 421 ASLAAKLGVTQCTIAHALEKTKYPSDSIYWKULEDKTHFSQCFADLIPANNHTDPIITTS 480
 DB 421 ASLAAKLGVTQCTIAHALEKTKYPSDSIYWKULEDKTHFSQCFADLIPANNHTDPIITTS 480
 QY 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
 DB 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
 QY 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
 DB 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
 QY 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKYNLNGQFRWISSQNNRIRNVELYRYI 660
 DB 601 ELANLVVVGDRRKESKDLSEKAEKMKFELIDKYNLNGQFRWISSQNNRIRNVELYRYI 660
 QY 661 CDTKGAFFQPALYEAFGITVVEAMTCGLPTPATCNGGPAEIIYVHGSGSPNIDPYHGDAA 720
 DB 661 CDTKGAFFQPALYEAFGITVVEAMTCGLPTPATCNGGPAEIIYVHGSGSPNIDPYHGDAA 720
 QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTTLGVYGFWMKXSNLEBR 780
 DB 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTTLGVYGFWMKXSNLEBR 780

```

Db      720 ELVDVFNCKIDQSHWDEISKGAMORIEEYTKWYISERLLNTLAVYGFWKHVTNIDRR 779
Qy      781 ESRRYLEMEFYALKYRKLAESVPLAE 806
        |||||
Db      780 ESRRYLEMEFYALKYRKPLAQSVPAVE 805

RESULT 4
AAB28141
ID      AAB28141 standard; protein; 805 AA.
AC      AAB28141;
DT      02-FEB-2001 (first entry)
DE      Sucrose synthase promoter protein #2.
KW      Promoter; eucalyptus; pine; gene transcription.
OS      Eucalyptus grandis.
XX      WO200058474-A1.
XX      05-OCT-2000.
XX      24-FEB-2000; 2000WO-NZ000018.
XX      25-MAR-1999; 99US-00276599.
XX      PR 30-JUL-1999; 99US-0146591P.
XX      PA (GENE-) GENESIS RES & DEV CORP LTD.
XX      PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX      PI Perera R, Rice SJ, Eagleton CK;
XX      WP1; 2000-647236/62.
XX      DR N-PSDB; AAC62807.
XX      PT Novel promoter sequences useful for modulating transcription of plant DNA
XX      PT sequences of interest and production of polypeptides.
XX      Claim 4; Page 70-72; 93pp; English.
XX      The present invention relates to promoter sequences from eucalyptus and
XX      CC pine. The present protein is encoded by one such promoter. The coding
XX      CC sequence for the present protein is useful for modulating the
XX      CC transcription of DNA sequences of interest. The sequences may also be
XX      CC used to tag or identify an organism or its reproductive material
XX      SQ Sequence 805 AA;

Query Match      85.8%; Score 3643.5; DB 3; Length 805;
Best Local Similarity 85.0%; Pred. No. 7,76-316;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

```

```

Db      241 GDTARVLEMTQLLDLLEAPDPTLEKFLDRVPMVNVVIMSPHGFADDDVLGYPTDG 300
Qy      301 GQVYVILDQVRALENEMTLRIKQGLNITPRILLITRLPDVAGTTGGRLKRYGTEHS 360
        |||||
Db      301 GQVYVILDQVRALKEEMLHRIKQGLDITPRILLITRLPDVAGTTGGRLKRYGTEHS 360
Qy      361 DILRVPRTEKGYRKWISREKWPYLETTYEDVAHEISKELHGTDLIIGNXSDGNIV 420
        |||||
Db      361 HILRVPRTEKGYRKWISREKWPYLETTYEDVAHEISKELHGTDLIIGNXSDGNIV 419
Qy      421 ASILAHKLVTOCTIAHALKTKYPPSDIYWKMLUEDCYHFSOCPADLFANNHTDFITTS 480
        |||||
Db      420 ASILAHKLVTOCTIAHALKTKYPPSDIYWKMLUEDCYHFSOCPADLFANNHTDFITTS 479
Qy      481 TPOELAGSKDVGQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYEPYTEKR 540
        |||||
Db      480 TPOELAGSKDVGQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYEPYTEKR 539
Qy      541 RLKHFPEIDDLTYKVENEBHLCVLDRNKKPIITMPRLDRVKNLTGLYEWCGKPKLR 600
        |||||
Db      540 RLKSFHPEIEELLFSDVENKEHLCVLDRNKKPIITMPRLDRVKNLTGLYEWCGKPKLR 599
Qy      601 ELANTLVVGGDRRKESKOLEKEMKMFELIDKYNLNGOPRWISSQMRIRNVELRYI 660
        |||||
Db      600 ELANTLVVGGDRRKESKOLEKEMKMFELIDKYNLNGOPRWISSQMRIRNVELRYI 659
Qy      661 CDTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCGGPAEIIYHSGSGFNIDPHYGDOAA 720
        |||||
Db      660 CDTKGAVVOPALYIAFGLTVVEAMTCGLPTFATCGGPAEIIYHSGSGFNIDPHYGDOAA 719
Qy      721 DIIVDFEKKCKQDSHWDKISQGLKRIIEEYTKWYISERLLITGYGFKHVSNLERR 780
        |||||
Db      720 ELVDVFNCKIDQSHWDEISKGAMORIEEYTKWYISERLLNTLAVYGFWKHVTNIDRR 779
Qy      781 ESRRYLEMEFYALKYRKLAESVPLAE 806
        |||||
Db      780 ESRRYLEMEFYALKYRKPLAQSVPAVE 805

RESULT 5
AAB80759
ID      AAB80759 standard; protein; 805 AA.
AC      AAB80759;
XX      26-MAR-2002 (first entry)
XX      DE Eucalyptus grandis promoter polypeptide #10.
XX      KW Eucalyptus grandis promoter polypeptide #10.
XX      KW Promoter; pine; leaf; root; flower; pollen; bud; meristem; xylogenesis;
XX      KW temporally regulated promoter; Pinus radiata; Eucalyptus grandis.
XX      OS Eucalyptus grandis.
XX      WO200198485-A1.
XX      27-DEC-2001.
XX      20-JUN-2001; 2001WO-NZ000115.
XX      20-JUN-2000; 2000US-00598401.
XX      PR 28-NOV-2000; 2000US-00724624.
XX      PA (GENE-) GENESIS RES & DEV CORP LTD.
XX      PA (FLET-) FLETCHER CHALLENGE FORESTS IND LTD.
XX      PI Perera R, Rice S, Eagleton C, Lasham A;
XX      WP1; 2002-114583/15.
XX      DR N-PSDB; ABK17072.
XX      PT Novel polynucleotide promoter sequences from pine and Eucalyptus useful
XX      PT for modifying expression of endogenous and/or heterologous

```


PT polynucleotides in transgenic plants.

XX Example 13; Page 86-88; 121pp; English.

CC The invention relates to isolated promoter sequences from *Pinus radiata*
 CC and *Eucalyptus grandis*, comprising a sequence chosen from leaf-, root-,
 CC flower-, pollen-, bud-, meristem-specific promoters or temporally
 CC regulated promoters such as xylogenesis-specific promoters. The promoter
 CC polypeptides and their related polynucleotides are useful in the
 CC production of genetic constructs, used for modifying gene expression in a
 CC target organism, in particular a plant. The method is useful for
 CC modifying expression of a polynucleotide that comprises an intron
 CC sequence, through removal of the intron sequence. The method is useful
 CC for modifying growth and development of plants, and cellular responses to
 CC external stimulus, such as environmental factors and disease pathogens.
 CC The sequences are useful in genome and physical mapping, in positional
 CC cloning of genes, in various assays to determine biological activity, to
 CC raise antibodies, to isolate corresponding interacting proteins and other
 CC compounds, and to quantitatively determine levels of interacting proteins
 CC or other compounds. Sequences AB080745-AA080764 represent *Pinus radiata*
 CC and *Eucalyptus grandis* polypeptides of the invention

XX Sequence 805 AA;

Query Match 85.8%; Score 3643.5; DB 5; Length 805;
 Best Local Similarity 85.0%; Pred. No. 7.7e-316;
 Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

QY 1 MAERALT RVHSRLERLDETLAHRNEILALSRIBEGKGILOHQIILEFEAIPENRK 60
 DB 1 MADRMLTRSHSLRERLDETLAHRNDIVAFLSRVEAKGIIQRIHQIFAEFAISESRA 60
 QY 61 KLANGAFPEVLASQGAIVLPWPVALAVPRPGWEYIVVNVHVALVEELTVAEYLHFEK 120
 DB 61 KLIDAFGEVLKSTGEAIVSPWVALAVPRPGWEHIVNVHVALVEELTVAEYLHFEK 120
 QY 121 ELVDSSNGNFVLELDFEFPNSFPPTLSKISGNGVEFLNHLISAKLPHDKESHMPLLE 180
 DB 121 ELADSSNGNFVLELDFEFPNTSFPPTLSKISGNGVEFLNHLISAKLPHDKESHMPLLE 180
 QY 181 FLRVHCHGKKNMMLNDRIQNALQVLRKAEYLGTLPEPTPCAEFEHRFOEIGLERGW 240
 DB 181 FLQVHCYKGNMNVNARIQNVSLQGVLRKAEYLTSLKPEPYSGFEHKEFGELGERGW 240
 QY 241 GPTARVLEMIQLLDLEATPCTIEKFLGRIPWVFNVVILTPHGYFQDNVLYGPPDG 300
 DB 241 GPTARVLEMIQLLDLEATPCTIEKFLGRIPWVFNVVILTPHGYFQDNVLYGPPDG 300
 QY 301 GGVVITLDQVRLAENEMLRIRKQOGNTPRIILITRLPDVVGTTCCGRLEKVVGTBHS 360
 DB 301 GGVVITLDQVRLAENEMLRIRKQOGNTPRIILITRLPDVVGTTCCGRLEKVVGTBHS 360
 QY 361 DILRVFPTREKGIVRKMSRFEKWPYLETYEDVAHEISKEIAGTPDILIGNXSDNVY 420
 DB 361 HILRVFPTREKGIVRKMSRFEKWPYLETYEDVAHEISKEIAGTPDILIGNXSDNVY 420
 QY 421 ASLLAHKLGVTCTTAHALEKTRPDSIYWKLEDEKTHFSCQFTADLPANMHTDPIITS 480
 DB 420 ASLLAHKLGVTCTTAHALEKTRPDSIYWKLEDEKTHFSCQFTADLPANMHTDPIITS 480
 QY 481 TPOEAGSKDVTGQYESHAFPLPGLYRVVHGIDVDFPNTVSPADMEIYFPYTEERK 540
 DB 481 TPOEAGSKDVTGQYESHAFPLPGLYRVVHGIDVDFPNTVSPADMEIYFPYTEERK 540
 QY 541 RLKHFPPEIEDLLYTVENEBHLCVLNDRNKPEILFTMPRLDVKVLTGLVENCGRNPKLR 600
 DB 540 RLKHFPPEIEDLLYTVENEBHLCVLNDRNKPEILFTMPRLDVKVLTGLVENCGRNPKLR 600
 QY 601 ELANLVVCGDRKRSKDLSEKAEKMKPELIDKTNLQGFWMISSQNNRINVELYRYT 660
 DB 601 ELANLVVCGDRKRSKDLSEKAEKMKPELIDKTNLQGFWMISSQNNRINVELYRYT 660
 QY 660 ELANLVVCGDRKRSKDLSEKAEKMKPELIDKTNLQGFWMISSQNNRINVELYRYT 660
 DB 660 ELANLVVCGDRKRSKDLSEKAEKMKPELIDKTNLQGFWMISSQNNRINVELYRYT 660
 QY 720 CDTKGFVQPALVEAFGLTVEAMTCGLPTFATCNGPABIIIVHSGSGNIDPYHGDQA 720

DB 660 CDTKGFVQPALVEAFGLTVEAMTCGLPTFATCNGPABIIIVHSGSGNIDPYHGDQA 719
 QY 721 DILVDFEKKCKDPDHPMKISQGGIKRTEKXTWYSEBRLITLGVYGFVHGVNLERK 780
 DB 720 ELVDVDFEKKCKDPDHPMKISQGGIKRTEKXTWYSEBRLITLGVYGFVHGVNLERK 779
 QY 781 ESRRYLEMFPYALKYRKLAEVPLAE 806
 DB 780 ESRRYLEMFPYALKYRKLAEVPLAE 805

RESULT 6

AB092501 standard; protein; 808 AA.

AC AB092501;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1712.

KW Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidner M;

DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.

XX Claim 5; SEQ ID NO 1712; 261pp + Sequence listing; English.

CC The invention relates to identifying target proteins (AB090790-AB094016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 808 AA;

Query Match 85.1%; Score 3614.5; DB 5; Length 808;
 Best Local Similarity 84.6%; Pred. No. 3e-313;
 Matches 681; Conservative 58; Mismatches 65; Indels 1; Gaps 1;

QY 2 AERALT RVHSRLERLDETLAHRNEILALSRIBEGKGILOHQIILEFEAIPENRK 61
 DB 4 AERVITRVHSGERLDAITVAQNEVFPALLSRVEAKGIIQRIHQIFAEFAIMPLETQK 63
 QY 62 LANGAFPEVLASQGAIVLPWPVALAVPRPGWEYIVVNVHVALVEELTVAEYLHFEK 121
 DB 64 LKGAFFELLSAQGAIVLPWPVALAVPRPGWEYIVVNVHVALVEELTVAEYLHFEK 123
 QY 122 LVDSSNGNFVLELDFEFPNSFPPTLSKISGNGVEFLNHLISAKLPHDKESHMPLLEF 181

Db 124 LVDGIKGNFTLEDPFPPNNAFPPTLNTKTYIGDQVEFLNRLHSLKLFHDKESLHPILKF 183
 Qy 182 LRVHCHGKNMMLNDRIQNLNALQHLRKAEYLGTLPPETPCAFEPHRFOEIGLERGMG 241
 Db 184 LRLSHGKNTLMNNRIQNLNTLQNLKRAEYLMELKPEITLYSFEHKGFOEIGLERGMG 243
 Qy 242 DTAENVLEMIOQLLDLEATDPTCTLEKPLGRIPMWPNVILTPHGYFQDNVLYGPDYGG 301
 Db 244 DTAENVLMIRLLDLLEAPDPTCTLENFGRIPMWPNVILSPHGYFQDNVLYGPDYGG 303
 Qy 302 QVVYILDOVRALENEMLRKIQGGLNTPRLLIITRLPDAGTTCGRLKGVYTESHD 361
 Db 304 QVVYILDOVRALETMLQRIKQGLNTPRLLIITRLPDAGTTCGRLKGVYTESHD 363
 Qy 362 ILRVFPEKGIIVRKWISRFKEMWPLYETVEDVAHEISKELGTPDILLGNKSDGNIVA 421
 Db 364 ILRVFPEKGIIVRKWISRFKEMWPLYETVEDVAHEISKELGTPDILLGNKSDGNIVA 422
 Qy 422 SLIAHKLGVTOCTIAHALEKTYPDSDIYWKLEDKYHPSQFTADLFAMNHTDFTITST 481
 Db 423 SLIAHKLGVTOCTIAHALEKTYPDSDIYWKLEDKYHPSQFTADLFAMNHTDFTITST 482
 Qy 482 FOEIGSDDTGQVESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 541
 Db 483 FOEIGSDDTGQVESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 542
 Qy 542 LKHFHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKPKLRE 601
 Db 543 LKHFHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKPKLRE 602
 Qy 602 LANLVVGGDRRKESKDLLEKAEKMKKMFELIDKYNLNGQFPMWISSQMRINNVLYRYIC 661
 Db 603 LANLVVGGDRRKESKDLLEKAEKMKKMFELIDKYNLNGQFPMWISSQMRINNVLYRYIC 662
 Qy 662 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGSPAEIIVHGKSGFHIDPYHGDQAD 721
 Db 663 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGSPAEIIVHGKSGFHIDPYHGDQAD 722
 Qy 722 ILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKLYSERLLTLTGVGFWKHSVNLBRE 781
 Db 723 SLADFTTCKNDPSHMDKISQGLKRIEKEYTWKLYSERLLTLTGVGFWKHSVNLBRE 782
 Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 Db 783 SRRYLEMFYALKYRKLAESVPLAEE 807
 RESULT 7
 ABB93562
 ID ABB93562 standard; protein; 808 AA.
 AC ABB93562;
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 2773.
 KW Herbicidal; plant; agriculture; herbicide.
 OS Arabidopsis thaliana.
 OS WO200210210-A2.
 PN 07-FEB-2002.
 PD 28-AUG-2001; 2001WO-EP009892.
 PF 28-AUG-2001; 2001WO-EP009892.
 PR 28-AUG-2001; 2001WO-EP009892.
 PA (FARB) BAYER AG.
 PI Tietjen K, Weidler M;
 XX

DR WPI; 2002-269010/31.
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.
 XX Claim 5; SEQ ID NO 2773; 261pp + Sequence Listing; English.
 PS
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 XX
 SQ Sequence 808 AA;
 Query Match 84.8%; Score 3603.5; DB 5; Length 808;
 Best Local Similarity 83.9%; Pred. No. 2.9e-312;
 Matches 675; Conservative 68; Mismatches 61; Indels 1; Gaps 1;
 Qy 2 AERALTTRVHSLREBLDETLLAHNRELLALLSRIGKKGKGIQHHQIIIEFPAIPEBNKK 61
 Db 4 AERMTTRVHSGREBLNETLVERNEVLALLSRVAKGKGIQHHQIIIEFPAIPEOTRKK 63
 Qy 62 LANGAFFVVKASQEAIVLPWVALAVPRGQWVEYIRVYNHAIIVSELTVAEVLHPEE 121
 Db 64 LEGGFEDLLSTQDAIVLPWVALAVPRGQWVEYIRVYNHAIIVSELTVAEVLHPEE 123
 Qy 122 LVDSSNGNFVLALDPEFPNPSFPRPLTSKISIGGVETFLNRLSAKLFHDKESMHPLEF 181
 Db 124 LVDGKGNFTLEDPFPPNNAFPPTLNTKTYIGDQVEFLNRLHSLKLFHDKESLHPILKF 183
 Qy 182 LRVHCHGKNMMLNDRIQNLNALQHLRKAEYLGTLPPETPCAFEPHRFOEIGLERGMG 241
 Db 184 LRLSHGKNTLMNNRIQNLNTLQNLKRAEYLMELKPEITLYSFEHKGFOEIGLERGMG 243
 Qy 242 DTAENVLEMIOQLLDLEATDPTCTLEKPLGRIPMWPNVILTPHGYFQDNVLYGPDYGG 301
 Db 244 DTAENVLMIRLLDLLEAPDPTCTLENFGRIPMWPNVILSPHGYFQDNVLYGPDYGG 303
 Qy 302 QVVYILDOVRALENEMLRKIQGGLNTPRLLIITRLPDAGTTCGRLKGVYTESHD 361
 Db 304 QVVYILDOVRALETMLQRIKQGLNTPRLLIITRLPDAGTTCGRLKGVYTESHD 363
 Qy 362 ILRVFPEKGIIVRKWISRFKEMWPLYETVEDVAHEISKELGTPDILLGNKSDGNIVA 421
 Db 364 ILRVFPEKGIIVRKWISRFKEMWPLYETVEDVAHEISKELGTPDILLGNKSDGNIVA 422
 Qy 422 SLIAHKLGVTOCTIAHALEKTYPDSDIYWKLEDKYHPSQFTADLFAMNHTDFTITST 481
 Db 423 SLIAHKLGVTOCTIAHALEKTYPDSDIYWKLEDKYHPSQFTADLFAMNHTDFTITST 482
 Qy 482 FOEIGSDDTGQVESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 541
 Db 483 FOEIGSDDTGQVESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 542
 Qy 542 LKHFHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKPKLRE 601
 Db 543 LKHFHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKPKLRE 602
 Qy 602 LANLVVGGDRRKESKDLLEKAEKMKKMFELIDKYNLNGQFPMWISSQMRINNVLYRYIC 661
 Db 603 LANLVVGGDRRKESKDLLEKAEKMKKMFELIDKYNLNGQFPMWISSQMRINNVLYRYIC 662
 Qy 662 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGSPAEIIVHGKSGFHIDPYHGDQAD 721
 Db 663 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGSPAEIIVHGKSGFHIDPYHGDQAD 722

Qy 722 ILVDFEKKCDPSHMDKISOGGLRIEKKYTWKYSRLTLTNGVGFWMKIVSNLERRE 781
 Db 723 TLADFETCKEDPSHMDKISOGGLRIEKKYTWKYSRLTLTNGVGFWMKIVSNLERRE 782
 Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 Db 783 ARRYLEMFYALKYRKLAESVPLAEE 807

RESULT 8

AAE28501 AAE28501 standard; protein, 816 AA.

AAE28501; AAE28501;

27-DEC-2002 (first entry)

Corn sucrose synthase (Sus1).

Sucrose synthase; shrunken-1; Sh1; constitutive sucrose synthase; Sus1; transgenic plant; plant breeding; grain breakage; grain strength; enzyme; cellulose; corn; chromosome 9.

Zea mays.

MO200267662-A1.

06-SEP-2002.

21-FEB-2002; 2002MO-US005137.

22-FEB-2001; 2001US-0270777B.

(PION-) PIONEER HI-BRED INT INC.

Dhuga KS, Helencjars TG, Niu X;

WPI; 2002-691625/74.

N-PSDB; AAD45852.

New polynucleotide and its encoded sucrose synthase, useful for modulating the level of sucrose synthase in transgenic plants (e.g. maize or soybean) to improve stalk length, reduce grain breakage, or improving plant or grain strength.

Example 9; Page 114-115, 125pp; English.

The invention relates to two sucrose synthases, shrunken-1 (Sh1) and constitutive sucrose synthase (Sus1) and their corresponding nucleic acids. The polynucleotide, or its encoded protein, is useful for modulating the level of sucrose synthase in a transgenic plant, increasing cellulose production in the stalk tissue of a transgenic plant, or increasing the concentration of cellulose in the tissues of a seed of a transgenic plant. This is particularly useful in plant (e.g. maize or soybean) breeding, especially for e.g. improving stalk length in maize, reducing grain breakage during combining, transport or movement in storage, or improving plant or grain strength. The present sequence is corn Sus1 protein. Corn Sus1 gene is located at chromosome 9

Sequence 816 AA;

Query Match 77.6%; Score 3295; DB 5; Length 816;

Best Local Similarity 76.8%; Pred. No. 1,1e-284;

Matches 618; Conservative 84; Mismatches 101; Indels 2; Gaps 2;

Qy 3 ERALTRVSLRLDELTLAHRNELIALSRIGKGILOHQQITLF-BAIPEENKX 61
 Db 7 DRVLSRLHSVERIGDSLSAHNDELVAFTLKNLKGKMLQPHQILAYNNALIPAAREK 66
 Qy 62 LANGAFEEVLAQSQAIVLPWVALAVPRGWEYIIVNTHALVEELTVAEYIHFKEE 121
 Db 67 LKDGAFEDVLRAAQAIVIPWVALAIRPRGWEYIVNTHALVEELTVAEYIHFKEE 126

Qy 122 LVDSNNGNFWLELDPEFPNNSFPRPTLSKISGNGVEFLNRLHSAKLFHDKESMPLLEF 181
 Db 127 LVESGPNNNFVLELDPEFPNNSFPRPTLSKISGNGVEFLNRLHSAKLFHDKESMPLLEF 186
 Qy 182 LRVHCHGKNNMLNDRIQNLALQVLRKAEYVGTLPETPCAFEEHRRFOELIGRWG 241
 Db 187 LRAHYKGMNTMMNDRIQNLALQVLRKAEYVGTLPETPCAFEEHRRFOELIGRWG 246
 Qy 242 DTAERVLEMIOQLDLLEATDPTLEKELGRIPWVFNVLITLPGYFQDQVVLGPDG 301
 Db 247 DCAKRAQSTHLLDLLEAPDPTLEKELGRIPWVFNVLITLPGYFQDQVVLGPDG 306
 Qy 302 QVVYILDQVRALENEMLRIRKQGLNITPRILITRLPDAVGTTCGRLEKVGTEHSD 361
 Db 307 QVVYILDQVRALENEMLRIRKQGLNITPRILITRLPDAVGTTCGRLEKVGTEHSD 366
 Qy 362 ILRVFPTKEGIVRKWISRFERKWPYLETYTEVAHEISKELHGTPLLIGNXSDGNIVA 421
 Db 367 ILRVFPTKEGIVRKWISRFERKWPYLETYTEVAHEISKELHGTPLLIGNXSDGNIVA 425
 Qy 422 SLNHLKGVTOCTTAHLEKTKYPSDLYMKKLEDKYHFSQFTADLFAMNHTPIITST 481
 Db 426 CLLAHKGVTHTCTTAHLEKTKYPSDLYMKKLEDKYHFSQFTADLFAMNHTPIITST 485
 Qy 482 FOEIASGKDTGVGOYESHPTAFTLPGLYRVVHGIDVFPDKNIVSGADMEIYFYTEEKR 541
 Db 486 FOEIASGKDTGVGOYESHPTAFTLPGLYRVVHGIDVFPDKNIVSGADMEIYFYTEEKR 545
 Qy 542 LKHFPETIEDLLYTKVENEELCYLANDRNPDLFTMPRLDRVKNLTGLVEMCGNPKLRE 601
 Db 546 LKSLHPEIEBELLYQTEWTEHFKVANDRNPDLFTMPRLDRVKNLTGLVEMCGNPKLRE 605
 Qy 602 LANLVVVGDDPKRSKSLDEKAEKMKAPLIDKYNLNGQFVWISSOMNRINVELYRIC 661
 Db 606 LANLVVVGDDPKRSKSLDEKAEKMKAPLIDKYNLNGQFVWISSOMNRINVELYRIC 665
 Qy 662 DTGAFVOPALYEAFTLVVAVMTGSLPTFATNGSPAEITVHGSGFNIPYHQDQAD 721
 Db 666 DTGAFVOPALYEAFTLVVAVMTGSLPTFATNGSPAEITVHGSGFNIPYHQDQAD 725
 Qy 722 ILVDFEKKCDPSHMDKISOGGLRIEKKYTWKYSRLTLTNGVGFWMKIVSNLERRE 781
 Db 726 TLADFETCKEDPSHMDKISOGGLRIEKKYTWKYSRLTLTNGVGFWMKIVSNLERRE 785
 Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 Db 786 TRRYLEMFYALKYRKLAESVPLAEE 810

RESULT 9

ADC07856 ADC07856 standard; protein, 816 AA.

ADC07856; ADC07856;

18-DEC-2003 (first entry)

Rice protein sequence Seq ID122 related to grain filling.

plant biotechnology; carbohydrate synthesis; carbohydrate metabolism; carbohydrate degradation; carbohydrate; plant grain; grain filling; corn; tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarcane; wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain; gene; ds; plant.

Oryza sativa.

MO2003000905-A2.

03-JAN-2003.

21-JUN-2002; 2002MO-IB002450.

PR 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 P1 Glazebrook J, Katagiri F, Kreps J, Provart N, Rieke D;
 XX
 DR MPI; 2003-229341/22.
 N-PSDB; ADC07855.
 PR New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.
 XX
 PS Claim 1; SEQ ID NO 122; 130bp; English.
 CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more
 CC desirable level using the present invention. The present sequence is the
 CC amino acid sequence of a rice protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPD
 CC at ftp.wipd.int/pub/publishedpet_sequences.
 XX
 SQ Sequence 816 AA;
 Query Match 77.1%; Score 3273; DB 7; Length 816;
 Best Local Similarity 76.1%; Pred. No. 9.9e-283;
 Matches 613; Conservative 91; Mismatches 99; Indels 2; Gaps 2;
 QY 3 BRALRVHSLRRLDETLAHHNELALSRLEGKGLQHQTILF-EAIPENRKK 61
 DB 7 DRLVSLRHSVRERIDSLSAHPNELVAVETRLVNLGKMLQHIIEVNNALISADREK 66
 QY 62 LANGAFEFYKASOBAIVLPWVALAVRPPGMEYIRVNVVALVVEELTVAEYHFKEE 121
 DB 67 LKDGAFEDVLRSAQEGIVSPWVALAIRPGEVMEYVNVSELAVELLTPVEYQFDEQ 126
 QY 122 LVYDGSNGNFVLELDFEPNSSFPRPTLSKSIINGVEFLNRHLAKLPHDKESMPLLEF 181
 DB 127 LVEEGTNNNFVLELDFEPNSSFPRPTLSKSIINGVQFLNRHLSPDKESMPLPLNF 186
 QY 182 LRVHCHKGKMMNLRIQNLNLQHLVLRKAEYVLTLPETPCAFERFQSIGLERMG 241
 DB 187 LRAHHYKGGTMMNLNRIKLSALQALRKAEEHLSGLSADYPSBFHHRFQELGLEKMG 246
 QY 242 DTAERVLLEWIOQLDLLEATDPTLEKFLGRIIPVENVVILTPHGYPADONVLAGPDDGG 301
 DB 247 DCAKASQETIHLDLLEAPDSTLEKFLGRIIPVENVVIMPHGYFAQANVLGTPDDGG 306
 QY 302 QVVVILDDVRALENEMTLRIKQGINTPRILITRLPDAVGTTCGRLERKVGTEHSD 361
 DB 307 QVVVILDDVRAMENEMTLRIKQGINTPRILITRLPDAVGTTCGRLERKVGTEHSD 366
 QY 362 ILRVPRTEKGIIVRKWISFEKWPVLETYEDVAHEISKELHGPDLITGXSGNIVA 421
 DB 367 ILRVPRTEKGIIVRKWISFEKWPVLETYEDVAHEISKELHGPDLITGXSGNIVA 425
 QY 422 SLIAHLGVTGCTIAHALEKTYPPDSDIYKKLEDPKHPSCQFTADLPFMMNTDPIITST 481
 DB 426 CLIAHMGVTHCTIAHALEKTYPPNSDLWKKFEEDHYHSCQFTDILNMAHADIIITST 485
 QY 482 FOEIASKDTVGQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYPPYTEKKR 541

DB 486 FOEIASKDTVGQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMSIYFPYSESRK 545
 QY 542 LKHFPRIEDLLYTKVNEEHLCTVANDRNKPIFTMPRLDRVKNLTGLVEMCGKPKLRE 601
 DB 546 LTSIHPRIEELLYSEVONNEHKPMLKDRNKPFIIFSMARLDRVKNLTGLVELYGNPNLOE 605
 QY 602 LANIVVGGRRKESKDLBEKAKENKPFELIDKYNLNGQPRWISSQNNRIRNVELYRYIC 661
 DB 606 LVNLVVVCGDHGNSKDEQDAEFKCMFDLEQYNLNGHIRMISAQNRVANGSELRYIYC 665
 QY 662 DTGAFVQPALYKFAFGTLVWEAMTCGLPTFATCNGSBAEITVHSGSPNIDPYGDDAAD 721
 DB 666 DTGAFVQPAFYKFAFGTLVWESMTGSLPTATAYGSPAEITVWVSGPHIDPYGDKASA 725
 QY 722 ILVDFPEKCKDPESHMDKISQGLKRIEKKYTWKIVSERLTLTGVGFMWGVNLSRRE 781
 DB 726 LLVFFPKCQDPESHMDKISQGLKRIEKKYTWKIVSERLTLTGVGFMWGVNLSRRE 785
 QY 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 DB 786 TRRYLEMLYALKYRTMASTVPLAVE 810
 RESULT 10
 ID ADC08209 standard; protein, 816 AA.
 XX
 AC ADC08209;
 DT 18-DEC-2003 (first entry)
 XX
 DE Rice protein sequence Seq ID514 related to grain filling.
 XX
 KW plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KW gene; ds; plant.
 XX
 OS Oryza sativa.
 XX
 PN WO200300905-A2.
 XX
 PD 03-JAN-2003.
 XX
 PE 21-JUN-2002; 2002WO-IB002450.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 P1 Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 P1 Glazebrook J, Katagiri F, Kreps J, Provart N, Rieke D;
 XX
 DR MPI; 2003-229341/22.
 N-PSDB; ADC08208.
 XX
 PT New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.
 XX
 PS Claim 34; SEQ ID NO 514; 130bp; English.
 CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,

CC wheat, and rice. The invention may be useful for the
CC improvement of protein, oil, starch, fibre and moisture content of the
CC cereal grains. In addition, carbohydrate levels may be modified to a more
CC desirable level using the present invention. The present sequence is the
CC amino acid sequence of a rice protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fp.wipo.int/pub/publishedpct_sequences.

Sequence 816 AA;

Query Match 77.1%; Score 3273; DB 7; Length 816;

Best Local Similarity 76.1%; Pred. No. 9.9e-283;
Matches 613; Conservative 91; Mismatches 99;

Matches 613; Conservative 91; Mismatches 99; Indels 2; Gaps 2

Qy	3	ERALTIVHSLRELUDETLAHNREITALLSRIEKGKGLIHOHOLIEF-BAIDENRK	61
Dh	7	DVALSRHLSVREKIGDLSAHNPELVAETRLVNLGKMLQHOIAEYNNALISEADREK	66
Qy	62	LAMGAFFELVKAQOEAVLVPWVALAVRPRGVMEXIRNNVALAVEELTAAEYLHKEE	121
Dh	67	LKOGAFEDVARSQOEGIVISPMVALAIRPRGMEYAVNNSELAVELLTYPEYLQKEQ	126
Qy	122	LVDGSSGNFVLELDEPPNSSPFRPTLSKISIGNGVEPLNRHLSAKLFHDYESMPLLEF	181
Dh	127	LVEEGTANNFVLELDEPPFNASFPFRPSLSKISIGNGVQFLNRHLSKLFHDKESMYPLNLF	186
Qy	182	LRYNCHGKGMMLNDIRIQUNLMQHLRAKAEYVLTLPPEPCAFEFHRPOEILBEGWG	241
Dh	187	LRAHYNGMMLNDIRISLSALQALRAEHEHLSGLSDAPYSEFNHRFOELBLEKGGW	246
Qy	242	DTAERVLEMIQLLDLLEATDPCTEKEFGIRIPMFNVNVLTPRPHYFADQNVLYGPTGG	301
Dh	247	DCAKRSETHLHLLDLLEAPDEBTEKEFGITIPMFNVNIMSPHGYFQANVLGPTGG	306
Qy	302	QVYVILDOVALENEMILRIKQGLNITERRLIITRLLPDAVGTTCGGRLEKVVYTESHD	361
Dh	307	QVYVILDOVAMENEMILRIKQGLNITERRLIITRLLPDAGTTCGGRLEKVGTEBTH	366
Qy	362	ILAVPRFTEGIYRKXKISREPKXPRVLEFYTDDVAHELKSELHGPRDLITGNXSQNVIA	422
Dh	367	ILAVPRFTEGIYRKXISRE-WPVLEFYTDDVAHELKSELQANPRDLITGNYSQNLVA	423
Qy	422	SLLAHKLGVTQCTIAHLKTKYPRDSDIYMKLDEKYHSPCOPTADLFAMNHTPTITST	481
Dh	426	CLAHKMGVHTCTIAHLKTKYPRNSDLWKKGFEDHYHPSQOPTDILAMNHDPITST	485
Qy	482	FOEIAGSKDVTGQYESHTAFTLPGLYVVHGDVDPKRNIVSPGADNEIIFYTEEKR	541
Dh	486	FOEIAGSKDVTGQYESHMAFTMGLYRVVHGDVDPKRNIVSPGADWSIIFYPSESRR	545
Qy	542	LKHFRHREIDLLYTKYVENEBHLCVUNDRKRPILTFMPLRDYRQNTGLVENCGNPKLRE	601
Dh	546	LTSLHREIEELLSYEVONNEHKMFLDNRRKPIIFSNARIDRYKNTGLVELYGNRPLOE	605
Qy	602	LANLVVVGSGRRKXESKDLEKAKMKMFELIDKYNLNGCFRWISSQMRIRINVELYRIC	661
Dh	606	LNVLVVVGSGRHNKPSDKXEQHAEFKMFPDLIDQYNLNGHIRNISQNMRRVANGELXYIC	665
Qy	662	DTKGAFLVOPALYEAFGILTVEAMTQCLPTFPATCNGGPAEIIYHGKSGNIDPYHGOAD	721
Dh	666	DTKGAFLVOPAFYEAFGILTVESMTQCLPTFPATAYGGPAEIIYNGVSGHIDPYOGDASA	725
Qy	722	LIVDPFEKCKDDSHMDKISQGLKRIIEKTYMKIYSEBLLTLNVGYFMGHVSNLEBRE	781
Dh	726	LIVDPFEKCKDDSHMTKISQGLQRIIEKTYMKIYSEBLLTLNVGYFMGHVSNLEBRE	785
Qy	782	SRRYLEMFYALKRYKLAESVPLAE	806
Dh	786	TRRYLEMLYALKRYTMASTVPLAVE	810

ADC68460
ID ADC68460 standard; protein; 815 AA

AC ADC68460;

DT 18-DEC-2003 (first entry)

Lolium perenne fructan biosynthesis protein SEQ ID NO:170.

KW Lolium perenne; Festuca arundinacea; lignin; fructan; tannin;

33 XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

XX
15 MAY 2003

XX		67 NOV 2000	0000Z	NYO	000000
XX		67 NOV 2000	0000Z	NYO	000000

XX
07-NOV-2001 3:00:11S-0337703P

XX
XX
(GENE-) GENETIC RES & DEV COM

PA (WRIG-) WRIGHTSON SEEDS LTD.
XY

PI Demmer U, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M,
PI Caulsbury KM Hall C

WPT : 2003-441544/41

DR N-PSDB; ADC68424.
xy

PT New polynucleotide e

PT canin in a plant.
vvy

PS Claim 18; SEQ ID NO 170; 240pp; English.
 YY

cc The present invention describes isolated

CC active in lignin, tructan and tannin biosynthetic pathways. Also described. (1) an isolated oligonucleotide probe or primer corresponding

SQ Sequence 815 AA

Query Match 77.0%; Score 3270.5; DB 7; Length 815;

Matches 613; Conservative 92; Mismatches 97; Indels 3; Gaps 3;

3 ERALTRVHSRLRDLDTLLAHRNEILALLSRIEGKGILQHHQIILEFE-ALPEENRKK 61

Db 7 DVLSTRHSVRERIGISLSAHNELNVAFTRLVNLGKMLQPHQIIAEVNTAIPAEAREK 66

QY 62 LANGAFEEYIKASQEAIVLPVWVALAVRPRGWMEYINVNHALVBEELYAEYLHNKEE 1221

Db 67 LKQGAEDYDLPRAAQEAIVISPMVALAIDRPRGWEYVAVNVSSELAVEELSYBETYLQEKQ 1266

QY 122 LVNDGSSNGNFVLELDEPPNSSFPRPTLSKISGNCVEFLNHLNLSAKLPHDKESHPLLEF 1811

Db 127 LVVGG-SNKDFVLELDEPPNNAFPPRPSLSKISGNGVQFLNHLNLSKLFHDKESYPLPLNF 1855

QY 182 LVNHCHKSGKMMNLNDRIQNLNLQVLAKAEYLGTLPRPTPCAFLPHRPOEGLERGWG 2411

Db 186 LRAHNYKMPMMNDVRSLSALOGALAKAEHHLSGLPADTPYSDFHHRFOELGEEKWG 245
Qy 242 DTAERVELMIOLLDLLEATDPCTLEKFLGRIPMVFENVVILTPHGYFAQDNVLGYPTDGG 301
Db 246 DCAKRAOETLHLLDLLEAPDPSTLEKFLGRIIPMFENVVILSPHGYFAQDNVLGYPTDGG 305
Qy 302 QVVYILDOVRALLENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSD 361
Db 306 QVVYILDOVRAMENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSD 365
Qy 362 ILRVFPERKGIIVRWKISRFEKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNVVA 421
Db 366 ILRVFPERKGIIVRWKISRFEKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNVVA 424
Qy 422 SLAAHLGVTOCTTIAHALEKTYRPSDIYWKKLDEKHFSCOPTADLEFAMNHTDPIITST 481
Db 425 CLIAHKMGVTHCTIAHALEKTYRPSDIYWKKLDEKHFSCOPTADLEFAMNHTDPIITST 484
Qy 482 FOEINAGSDTVGQYESHSTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTYEEKR 541
Db 485 FOEINAGSDTVGQYESHSTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTYEEKR 544
Qy 542 LKHFRPELDELTYTVENEENHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMGKNPKLRE 601
Db 545 LTSIHPLEIELLYSDVNDENHFKVLDNRNKPILFTMPRLDRVKNLTGLVEMGKNPKLRE 604
Qy 602 LANLVVGGDRRKESKDLLEKAEKMKMPELIDKYNLNGQFWMISSQMRIRNVELYRYIC 661
Db 605 LNVNLVVGCDHGNPNKDKSEQAEFKKMDLIHQYNLNGHIRKISQMRIRNVELYRYIC 664
Qy 662 DTGAFVOPALYEAFTLVEMAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQAD 721
Db 665 DTGAFVOPALYEAFTLVEMAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQAD 724
Qy 722 ILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKLYSRLLLTLCGVGPMKHSVLSERE 781
Db 725 ILVDFEKKCKDPSHMDKISOGGLKRIEKKYTWKLYSRLLLTLCGVGPMKHSVLSERE 784
Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
Db 785 TRRYLEMLYALKYRKLAESVPLAEE 809

RESULT 12
AAE28500
ID AAE28500 standard; protein; 802 AA.
XX AAE28500;
AC AAE28500;
XX
DT 27-DEC-2002 (first entry)
XX
DE Corn sucrose synthase shrunken-1 (Sh1) protein.
XX
KW Sucrose synthase; shrunken-1; Sh1; constitutive sucrose synthase; Sus1;
KW transgenic plant; plant breeding; grain breakage; grain strength; enzyme;
KW cellulose; corn; chromosome 9.
XX
OS Zea mays.
XX
PN W0200267662-A1.
XX
PD 06-SEP-2002.
XX
PE 21-FEB-2002; 2002MO-US005137.
XX
PR 22-FEB-2001; 2001US-0270777P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhuga KS, Helentjaris TG, Niu X;
XX WPI; 2002-691625/74.
DR N-PSDB; AAD45851.

XX New polymucleotide and its encoded sucrose synthase, useful for
PT modulating the level of sucrose synthase in transgenic plants (e.g. maize
PT or soybean) to improve stalk length, reduce grain breakage, or improving
PT plant or grain strength.
XX
PS Example 9; Page 108-110; 125pp; English.
XX
CC The invention relates to two sucrose synthases, shrunken-1 (Sh1) and
CC constitutive sucrose synthase (Sus1) and their corresponding nucleic
CC acids. The polymucleotide, or its encoded protein, is useful for
CC modulating the level of sucrose synthase in a transgenic plant,
CC increasing cellulose production in the stalk tissue of a transgenic
CC plant, or increasing the concentration of cellulose in the tissues of a
CC seed of a transgenic plant. This is particularly useful in plant (e.g.
CC maize or soybean) breeding, especially for e.g. improving stalk length in
CC maize, reducing grain breakage during combining, transport or movement
CC into storage, or improving plant or grain strength. The present sequence
CC is corn Sh1 protein. Corn Sh1 gene is located at chromosome 9
XX
SQ Sequence 802 AA;
XX
Query Match 77.0%; Score 3268.5; DB 5; Length 802;
Best Local Similarity 76.2%; Pred. No. 2.4e-282;
Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;
Qy 6 LTRVHSRERLDETTLAHRNEILLRSIEGKGIIQHQIILEFPAIPENRKLKANG 65
Db 5 LTRHSRERLGAFFSSHPNELIALFSRYVHQGGMQLRQHLAEFDLPDSDEKKA-- 62
Qy 66 AFEEVLKASQEAITYLPPWVALAVRPPQWEYIRVNVYHALVVELTYAEIYHFEELVDG 125
Db 63 PFEDILPAAOEAIALPPWVALAIRPPQWMDYIRVNVSELAVERLSVEYLAFFKQVLVDG 122
Qy 126 SSNNFPLELDFEPPNSPFRPTLSKISGVVERLNLHLSAKLPHDKSMKPLLEFLRVH 185
Db 123 QSNNSFVLELDFEPPNSPFRPTLSKISGVVERLNLHLSAKLPHDKSMKPLLEFLRVH 182
Qy 186 CHKSKMMLNDRIQNLNALQHVLRKAEVYLTLPETPCAEFFHRFOELGEEKMGDTAE 245
Db 183 NYKGTMMNDRIQNLNALQHVLRKAEVYLTLPETPCAEFFHRFOELGEEKMGDTAK 242
Qy 246 RYLEMIOLLDLLEATDPCTLEKFLGRIIPMFENVVILTPHGYFAQDNVLGYPTDGGQVY 305
Db 243 RYLDLTHLLDLLEAPDPSTLEKFLGRIIPMFENVVILSPHGYFAQDNVLGYPTDGGQVY 302
Qy 306 ILDOVRALLENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSDIRV 365
Db 303 ILDOVRALLENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSDIRV 362
Qy 366 PFRTKGIIVRWKISRFEKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNVVA 425
Db 363 PFRTKGIIVRWKISRFEKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNVVA 421
Qy 426 HKLGVTCTIAHALEKTYRPSDIYWKKLDEKHFSCOPTADLEFAMNHTDPIITSTPEI 485
Db 422 HKLGVTCTIAHALEKTYRPSDIYWKKLDEKHFSCOPTADLEFAMNHTDPIITSTPEI 481
Qy 486 AGSKDTGQYESHSTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTYEEKRLKGF 545
Db 482 AGSKDTGQYESHSTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPTYEEKRLKGF 541
Qy 546 HPEIEDLLTYTVENEENHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMGKNPKLRELANT 605
Db 542 HPEIEDLLTYTVENEENHLCVLDNRNKPILFTMPRLDRVKNLTGLVEMGKNPKLRELANT 601
Qy 606 VVVGDRRKESKDLLEKAEKMKMPELIDKYNLNGQFWMISSQMRIRNVELYRYICDTKG 665
Db 602 VIVAGDHGKESKDLLEKAEKMKMPELIDKYNLNGQFWMISSQMRIRNVELYRYICDTKG 661
Qy 666 AFVOPALYEAFTLVEMAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQADILVD 725
Db 662 AFVOPALYEAFTLVEMAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQADILVD 721

PF 01-MAR-1999; 99JP-00052102.
 XX 01-MAR-1999; 99JP-00052102.
 XX (MITA) MITSUI CHEM INC.
 PA WPI; 2000-675173/66.
 DR N-PSDB; AAC66090.
 XX Novel method for the reinforcement of cold resistance in a plant
 PT comprising introducing a vector encoding an enzyme that decomposes
 PT sucrose into the plant.
 XX
 XX Disclosure; Page 8-11; 22pp; Japanese.
 XX
 XX This invention relates to a method for the reinforcement of cold
 CC resistance in a plant, comprising introducing an expression vector having
 CC a DNA encoding an enzyme for decomposing sucrose connected downstream to
 CC a promoter expressible in the plant, and expressing the enzyme in the
 CC plant body. Included in the invention are an expression vector used in
 CC the method; a transformed plant carrying the expression vector; and a
 CC transformed rice carrying the above expression vector. The method is used
 CC for reinforcing cold resistance in a plant. The present sequence
 CC represents a sucrose synthase protein used in the method
 XX
 XX Sequence 808 AA;

Query Match 76.7%; Score 3259.5; DB 3; Length 808;
 Best Local Similarity 76.2%; Pred. No. 1.6e-281;
 Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;

6 LTRVSLREBLDETLAHRNELIALSLRLEGKGILOHIOIIEFEALPEENRKLKANG 65
 5 LARLSLREBLGATSSHPNELIALFSRYNGKGMLOHQLAEFDALIEADKEKVA-- 62
 66 AFEVLKASQEAIVPWPVALAVRPPGWEIIRVNVHAYVEELTVAEYLHFKELVDG 125
 63 PREDILRAQEAIVPWPVALAIRPPGWDYIRVNVSELAELSEVSEYLAIFKQLVDG 122
 126 SNGNFVLELDEPPNSPFPPTLSKISGNGVEFNRHLSAKLFDKESMPLEFLRVH 185
 123 HTNSFVLELDEPPNSPFPPTLSKISGNGVEFNRHLSAKLFDKESMPLEFLRVH 182
 186 CHKGKMMMLNDRIQNLMLNLOHVLRAEYLGTLPEPTCAEFERHFOEIGERNGDTAE 245
 183 NHKGTMMMLNDRIQNLMLNLOHVLRAEYLGTLPEPTCAEFERHFOEIGERNGDTAE 242
 246 RVLNMIQILLLELXTPCTLEKFGRIPMVNVVITLPHGFPADNVLGYDTGGQVYV 305
 243 RVLNMIQILLLELXTPCTLEKFGRIPMVNVVITLPHGFPADNVLGYDTGGQVYV 302
 306 ILDOVRALENEMMLIKOQGLNITPRILITLRLPDVAGTTCGRLKYYGTEHSDIRV 365
 303 ILDOVRALENEMMLIKOQGLNITPRILITLRLPDVAGTTCGRLKYYGTEHSDIRV 362
 366 PERTEKGIYRKMISSFEKVPYLETYETDVAHEISKEELHGTBDLIIGNXSDGNIYASDLA 425
 363 PERSENGILRKMISSFEKVPYLETYETDVAHEISKEELHGTBDLIIGNXSDGNIYASDLA 421
 426 HGLSTVCTIAALAEKTKYPPSDIYWKLEDKYHNSCOFTADLPAMNHTDFIITSTPOEI 485
 422 HGLSTVCTIAALAEKTKYPPSDIYWKLEDKYHNSCOFTADLPAMNHTDFIITSTPOEI 481
 486 AGSKDTVGQESHHTAFTLPGLVVHGVDPDPKNNIYSPGADMEIYPPYTEKSKRLKHF 545
 482 AGSKDTVGQESHHTAFTLPGLVVHGVDPDPKNNIYSPGADMEIYPPYTEKSKRLKHF 541
 546 HPEIEDLLYTKVNEEHLCLVINDRKPILFTMPRLDRVNLTLGVLECGCKNKLRELANL 605
 542 HPEIEELLYSEVENDEHKFVLKDKKKPIFFSMARLDVKNMTGLVEMYGKANHLADLANL 601
 606 VVVGDRRESKDLBEKAKMKKFFLIDKYNLNGCPRTSSQWNRIRVVELYRYICDTKG 665

DB 602 VIYCGDHNGSKRDEQAEFKOMGLIDYKLRKIHIMISQNNRVNNGELRYRICDTKG 661
 QY 666 AFVOPALYEAFTLVEMAMTGLPTFATCNGSPAEIIVHGKSGFNIDPYHGDADILVD 725
 DB 662 VFQAPAFYEAFTLVEMAMTGLPTFATCNGSPAEIIVHGKSGFNIDPYHGDADILVD 721
 QY 726 FFEKCKDPSHWDKISGGGLKRIEKKYTWKYSRLTLTGVYGFVKVNSLERRESRRY 785
 DB 722 FFEKCKDPSHWDKISGGGLKRIEKKYTWKYSRLTLTGVYGFVKVNSLERRESRRY 781
 QY 786 LEMFYALKYRKLAEVPLA 804
 DB 782 IEMFYALKYRKLAEVPLA 800

RESULT 15
 AD68459
 ID AD68459 standard; protein, 808 AA.
 AC AD68459;
 DT 18-DEC-2003 (first entry)
 XX
 XX *Lotium perenne* fructan biosynthesis protein SEQ ID NO:169.
 DE *Lotium perenne*, *Festuca arundinacea*; lignin; fructan; tannin;
 KW biosynthetic pathway; plant.
 XX
 XX *Lotium perenne*.
 XX WO2003040306-A2.
 XX
 XX 15-MAY-2003.
 PD 07-NOV-2002; 2002WO-NZ000239.
 XX
 XX 07-NOV-2001; 2001US-0337703P.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (WRIG-) WRIGHTSON SEEDS LTD.
 XX
 PI Demmer J, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M,
 PI Saulsbury KM, Hall C;
 XX
 XX WPI; 2003-441544/41.
 DR N-PSDB; AD68423.
 XX
 XX New polynucleotide encoding polypeptides from *Lotium perenne* or *Festuca*
 PT *arundinacea*, useful for modulating the biosynthesis of lignin, fructan or
 PT tannin in a plant.
 XX
 XX
 PS Claim 18; SEQ ID NO 169; 240pp; English.
 CC The present invention describes isolated polynucleotides (i) encoding
 CC proteins (ii) from *Lotium perenne* and *Festuca arundinacea* which are
 CC active in lignin, fructan and tannin biosynthetic pathways. Also
 CC described: (1) an isolated oligonucleotide probe or primer comprising at
 CC least 10 contiguous residues complementary to 10 contiguous residues of
 CC (i); (2) a kit comprising the oligonucleotide probe or primer; (3) a
 CC genetic construct comprising (i); (4) a transgenic plant cell comprising
 CC the genetic construct of (3); (5) a plant or its seed, fruit or progeny
 CC comprising the transgenic plant cell of (4); (6) modulating one or more
 CC of the lignin, fructan or tannin compositions of a plant; (7) producing a
 CC plant having one or more of the lignin, fructan or tannin compositions;
 CC and (8) modifying the activity of (ii) involved in a lignin, fructan or
 CC tannin biosynthetic pathway in a plant. (i) can be used for modulating
 CC the biosynthesis of lignin, fructan or tannin in a plant. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 XX Sequence 808 AA;

Query Match 76.6%; Score 3254.5; DB 7; Length 808;
 Best Local Similarity 76.2%; Pred. No. 4.4e-281;

		Matches	609;	Conservative	92;	Mismatches	95;	Indels	3;	Gaps	2;
Qy	6	LTRVSLREERDETLAARNELALLSRLEGKGLLOHQHIIIEFAIPEENRKLANG	65								
Db	5	LTRLHSLEERLGATSSHPNELALPSKYVHGKGMLOHQHLLTEFEALFEADKERYA--	62								
Qy	66	AFPEVLKASOEALVLPVVALAVRPPGVWEYIRVNVVALVVEELTVAEYLHFKEELVDG	125								
Db	63	PFEDILRAAQEALIVLPVVALAIRPPGVMDYIRVNVSELAVEELTVSEYLAFKQLVDE	122								
Qy	126	SSNGNFVLELDEPPNSSFPPPTLSKSIQNGVEFLNRHLSAKLFHDKESMHPLEFLRVH	185								
Db	123	HASSKFVLELDEPPNSSFPPPSMSKSIQNGVQFLNRHLSKLFQDKESLYPLNFKAH	182								
Qy	186	CHKGNNMLNDRICQNLNLOHVLRAKAEELGLPPEPCAEFEHRFOEIGLERGWDTAE	245								
Db	183	NHKGTTMMLNDRIOGLRGLQSALRAEEYLTSPEDTPSSEFNHRFOEIGLEKMGWDTAK	242								
Qy	246	RVLWMIQLLLDLLEATDPCCTLEKFLGRIPMVFNVIILTPHGYFAODNVLGYPDGTGOVYV	305								
Db	243	RVQDTIHLLDLLEAPDPALEKFLGTIPMFMNVILSPHGYFAQSNVLGYPDGTGOVYV	302								
Qy	306	ILDQVRALENEMLRIRKQGLNITPRILITRLPDAVGTTCGRLEKVGTEHSDILRV	365								
Db	303	ILDQVRALENEMLRIRKQGLNITPRILITRLPDAVGTTCGRLEKVGTEHSDILRV	362								
Qy	366	PPRTKGIVRKVISPEKWPYLETYTEDVAHEISKELHGTPLDITIGNXSDGNIVASLLA	425								
Db	363	PPRTKGIVRKVISPEKWPYLETYTEDVAHEISKELHGTPLDITIGNXSDGNIVASLLA	421								
Qy	426	HKLGVTOCTIAHALKTKYPSDSIYWKKLEDKYHSCQFTADLFAMNHTDFIITSTPOEI	485								
Db	422	HKLGVTOCTIAHALKTKYPSDSIYLDKFDSDQYHSCQFTADLIAMNHTDFIITSTPOEI	481								
Qy	486	AGSKDTVGQYESHTAFTLPGLYRVVHGIDVPDKFNIVSPGADMEIYPPYTEKRRLLKHF	545								
Db	482	AGSKDVSQYESHIAFTLPDLRYRVVHGIDVPDKFNIVSPGADMTVYFPYTEDKRLTAF	541								
Qy	546	HPEIEDLYTKYNEEHLCLVDRNKPILFTMPRLDRVKNLTGLVEMCGKPKLRELNL	605								
Db	542	HPEIEELLYSDVENSEHKEFVLKDKNKPILFTSMARLDRVKNMTGLVEMFGKNAHLKDNL	601								
Qy	606	VVVGDRRKESKDLREKAKMKQFELIDKYNLNGQFRMISQMRIRNVELRYICDTKG	665								
Db	602	VIVAGDHGKESKDRREQAEPKMYSLIEBYKLEGHIRMSAQMRVRYNAELRYICDTKG	661								
Qy	666	AFVQPALYEAFLTYVEAMTCGLPTFATCNGGPABEIIVHGKSGFNIIDPYHGDQADILVD	725								
Db	662	AFVQPAFYEAFLTYVEAMTCGLPTATCHGSPAIIIVNGVSGLHIDPYHSDKADILVN	721								
Qy	726	FPEKCKDPSSHWDKISQGLKRIEKYWKIYSERLTLTGYYGFWKAVSNLERRESRY	785								
Db	722	FPEKSTADPTVYDKMSEGLKRIYEWKLYSERLMTLTGYGFWKAVSNLERRETRY	781								
Qy	786	LEMFYALKYRKLAESVPLA	804								
Db	782	LEMFYALKYRSILAASVPLA	800								

Search completed: June 2, 2004, 14:51:05
Job time : 66 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:49:19 ; Search time 21 Seconds
(without alignments)
3691.922 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247

Sequence: 1 MAERALTFRVHSLRERLDETL.....EMFYALKYRKLAESVPLAEE 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3669.5	86.4	806	S31479	sucrose synthase (
2	3614.5	85.1	808	T49233	sucrose synthase-1
3	3501.5	82.4	805	YUP05	sucrose synthase (
4	3400.5	80.1	808	S37560	sucrose synthase (
5	3268.5	77.0	802	1 YUZMS	sucrose synthase (
6	3264	76.9	816	S31939	sucrose synthase (
7	3259.5	76.7	808	S23543	sucrose synthase (
8	3238.5	76.3	801	T14338	sucrose synthase (
9	3222	75.9	816	S32451	sucrose synthase (
10	3209	75.6	807	S29242	sucrose synthase (
11	2985	70.3	809	2 T06497	sucrose synthase (
12	2962	69.7	809	2 B85029	probable sucrose s
13	2958	69.6	822	S71493	sucrose synthase (
14	2892	68.1	808	2 T01420	sucrose synthase (
15	2815	66.3	804	1 YUMU	sucrose synthase (
16	2479.5	58.4	946	2 S24966	sucrose synthase (
17	2312.5	54.5	982	2 C96760	probable sucrose s
18	1747	41.1	806	2 A17428	sucrose synthase (
19	1435.5	33.8	805	2 A11938	sucrose synthase (
20	941	22.2	235	2 J10281	sucrose synthase (
21	883	20.8	212	2 J10280	sucrose synthase (
22	803.5	18.9	204	2 S22535	sucrose synthase (
23	756	17.8	179	2 S22537	sucrose synthase (
24	661	15.6	218	2 S21311	sucrose synthase (
25	621	14.6	127	2 A29484	sucrose synthase (
26	494.5	11.6	720	2 S75935	hypothetical prote
27	437	10.3	1059	2 T12195	sucrose-phosphate
28	436	10.3	1068	1 J01329	sucrose-phosphate
29	435.5	10.3	1083	2 T04062	sucrose-phosphate

30	431	10.1	1056	2 J02277	sucrose-phosphate
31	421.5	9.9	1045	2 S55253	sucrose-phosphate
32	418.5	9.9	1081	2 T09837	sucrose-phosphate
33	413.5	9.7	1053	2 S34172	sucrose-phosphate
34	407	9.6	1057	2 S72648	sucrose-phosphate
35	404	9.5	1047	2 T51800	sucrose-phosphate
36	400	9.4	1084	2 T04103	sucrose-phosphate
37	394	9.3	1054	2 T09833	sucrose-phosphate
38	387.5	9.1	1049	2 J04783	sucrose-phosphate
39	386	9.1	1064	2 E86182	hypothetical prote
40	294	6.9	100	2 T09856	sucrose synthase (
41	256	6.0	67	2 T12251	sucrose synthase (
42	248	5.8	454	2 AG2657	glycosyltransferase
43	248	5.8	454	2 D97439	hypothetical prote
44	206.5	4.9	501	2 T01981	sucrose-phosphate
45	197.5	4.7	341	2 S72649	sucrose-phosphate

ALIGNMENTS

RESULT 1

S31479

sucrose synthase (EC 2.4.1.13) - fava bean

C.Species: Vicia faba (fava bean)

C.Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999

C.Accession: S31479

R.Heim, U.; Weber, H.; Wobus, U.

submitted to the EMBL Data Library, December 1992

A.Description: Sucrose synthase expression patterns in developing seeds of Vicia faba in

A.Reference number: S31479

A.Accession: S31479

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-806 <HEI>

A.Cross-references: EMBL:X69773; NID:G22037; PIDN:CAA49428.1; PID:G22038

C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology

C.Keywords: glycosyltransferase; hexosyltransferase

F:273-753/Domain: sucrose/sucrose-phosphate synthase homology <SSFS>

Query Match	86.4%; Score 3669.5; DB 2; Length 806;
Best Local Similarity	86.2%; Pred. No. 2.8e-235;
Matches	695; Conservative 52; Mismatches 58; Indels 1; Gaps 1;
QY	1 MAERALTFRVHSLRERLDETLAHRNEIALLSRIBGKGGILQHQTLEFBAIPEERK 60
DB	1 MATERTVHSLRERLDETLTANNREIALLSRIBGKGGILQHQTLEFBAIPEERK 60
QY	61 KLANGAFPEVLKASQEAIVLPWVALAVRPGWERYLRVWVHALVVEELVAEYLHPRKE 120
DB	61 KLTDGAFGEVLRSTQEAIVLPWVALAVRPGWERYLRVWVHALVVEELVAEYLHPRKE 120
QY	121 ELVDSGSGNFVLLEIDPEFPNSFPPTLSKISGVGFELNRHLSAKLFHDKESMPLLE 180
DB	121 ELVDSGSGNFVLLEIDPEFPPTASFPPTLSKISGVGFELNRHLSAKLFHDKESMPLLE 180
QY	181 FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYLGTLPPETPCAFEFHRQOIELENGW 240
DB	181 FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYLGTLPPETPCAFEFHRQOIELENGW 240
QY	241 GDTARVLEMLQLLDLLEATDPTLEKFLRIPVFNVLITPHGYFAQONNVGYPTG 300
DB	241 GDSARVLESTQLLDLLEADPTCTLEFLRIPVFNVLITPHGYFAQONNVGYPTG 300
QY	301 GQVYVILDOVALLENEMLRIRKQOGLNTPRILITRLLPDAVGTGQRLKXYGTSHS 360
DB	301 GQVYVILDOVALLENEMLRIRKQOGLDIVRILITRLLPDAVGTGQRLKXYGTSHS 360
QY	361 DILRVFTEKGIYRKATSRFEKWPYLETTEDEVNAHSKELGTPPLIIGNSDGIYV 420
DB	361 HILRVFTEKGIYRKATSRFEKWPYLETTEDEVNAHSELGTPPLIIGNSDGIYV 420
QY	421 ASLAAKLGIVGQCTAAHLEKTKYRDSDIYWKLEDKXKHFSCQFTADLFANNHTDFTTS 480

```
Db 420 ASLAHKLGVQCTTAHLEKTKYPESDIYWKKEEKKHFSOFTADLFAMNHTDFIITS 479
Qy 481 TFOELAGSKDTGVQYESHSTAFTLPGLYRVVGHIDVDPKFNIVPSGADMEIYFPYTEERK 540
Db 480 TFOELAGSKDTGVQYESHSTAFTLPGLYRVVGHIDVDPKFNIVPSGADMEIYFPYTEERK 539
Qy 541 RLKHFHPEIEDLLYTKVNEEHLCVLNDNRNPIFTPMRLDRVKNLTGLVEMCGKNPGLR 600
Db 540 RLTSFYPPIEBLLYSTVENEHLICVLKDRSKPIIFTMARLDRVKNITGLVEMYGKNAKL 599
Qy 601 ELANLVVVGSDRRKESKDLKEAKEMKMFELIDKYNLNGQFRWISSQMRIRANVELYRYI 660
Db 600 ELVNLVVVAGDRRKESKDLKEAKEMKMYELIETKJLNGQFRWISSQMRIRANVELYRYI 659
Qy 661 CDTGAFVQPALYEAFTGLTVEAMTCGPTFATCNGGPAEIIYHKGSGFNIDPYHGDOAA 720
Db 660 CDTGAFVQPALYEAFTGLTVEAMTCGPTFATCNGGPAEIIYHKGSGFNIDPYHGDOAA 719
Qy 721 DILVDFEKKCKDDPSHMWKISQGLKRIEKKYTWKIYSERLTLTGYYGFMKHYVSNLERR 780
Db 720 DILVDFEKKCKDDPSHMWKISQGLKRIEKKYTWKIYSERLTLTGYYGFMKHYVSNLERR 779
Qy 781 ESRRYLEMFPALYKRYKLAESVPLAE 806
Db 780 ESRRYLEMFPALYKRYKLAESVPLAE 805
```

RESULT 2

```
T49233
sucrose synthase-like protein - Arabidopsis thaliana
N:Alternate names: protein F7K15.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C:Accession: T49233
R:Obmaler: B.; Oetemaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lem
submitted to the Protein Sequence Database, April 2000
A:Accession: T49233
A:Reference number: 225019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-808 <OBS>
A:Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7K15.40
A:Experimental source: cultivar Columbia; BAC clone F7K15
C:Genetics:
A:Gene: ATSP:F7K15.40
A:Map position: 3
A:Introns: 35/2; 128/2; 192/3; 304/3; 336/3; 394/3; 433/3; 489/2; 564/2; 752/2; 798/3
C:Superfamily: sucrose synthase; sucrose-phosphate synthase homology
```

```
Query Match 85.1%; Score 3614.5; DB 2; Length 808;
Best Local Similarity 84.6%; Pred. No. 1.2e-231;
Matches 681; Conservative 58; Mismatches 65; Indels 1; Gaps 1;

Qy 2 AERALLTVHSLRELRDELTLAHRNEILLASRIEKGKGILOHQIILEFEAIPENRKK 61
Db 4 AERVITRVHSGERLDATLVAQKNEVPALLSRVEAKGKGILOHQIILEFEAMPLETOKK 63
Qy 62 LANGAFPEVLKASQDAIVLPWPVALAVRPFGWYIRVNVNVALVVEELVAEYLHFKEE 121
Db 64 LKGGAFPEFELKASQDAIVLPWPVALAVRPFGWYIRVNVNVALVVEELVAEYLHFKEE 123
Qy 122 LVDGSSNGNFVLELDFEPNNSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMHPLEF 181
Db 124 LVDGKNGNFVLELDFEPNNAFPRTLTKYIGDVEFLNRHLSAKLPHDKESMHPLEF 183
Qy 182 LRVHCHKRKNMNLDRIQNLALQVLRKAEYLGTLPPETPCAEFHRFOEIGLERGNG 241
Db 184 LRVHCHKRKNMNLDRIQNLALQVLRKAEYLMELKPELTYSEFEKHFQEGIGLERGNG 243
Qy 242 DTAEEVLEMIQILLDLLEKTPCTLEKPLGRIPMFPNVVILTPHGYFPQDNVLTGIPDDG 301
Db 244 DTAEEVLEMIQILLDLLEKTPCTLEKPLGRIPMFPNVVILTPHGYFPQDNVLTGIPDDG 303
```

```
Qy 302 QVVYIILDQVRALENEMELRIKQGLNITPRLITRLLPDAVGTTCQRLKRYGTGESHSD 361
Db 304 QVVYIILDQVRALENEMELRIKQGLNITPRLITRLLPDAAGTTCQRLKRYGTGESHSD 363
Qy 362 ILRVFPTTEGIVYRKVTSREKWPVLETYTEDVAHISKEHLNHTPDLITGNSSDGNIVA 421
Db 364 ILRVFPTTEGIVYRKVTSREKWPVLETYTEDVAHISKEHLNHTPDLITGNSSDGNIVA 422
Qy 422 SLAHKLGTVQCTTAHLEKTKYPESDIYWKKEEKKHFSOFTADLFAMNHTDFIITS 481
Db 423 SLAHKLGTVQCTTAHLEKTKYPESDIYWKKEEKKHFSOFTADLFAMNHTDFIITS 482
Qy 482 TFOELAGSKDTGVQYESHSTAFTLPGLYRVVGHIDVDPKFNIVPSGADMEIYFPYTEERK 541
Db 483 TFOELAGSKDTGVQYESHSTAFTLPGLYRVVGHIDVDPKFNIVPSGADMEIYFPYTEERK 542
Qy 542 LKHFHPEIEDLLYTKVNEEHLCVLNDNRNPIFTPMRLDRVKNLTGLVEMCGKNPGLR 601
Db 543 LKHFHPEIEDLLYTKVNEEHLCVLNDNRNPIFTPMRLDRVKNLTGLVEMCGKNPGLR 602
Qy 602 LANLVVVGSDRRKESKDLKEAKEMKMFELIDKYNLNGQFRWISSQMRIRANVELYRYI 661
Db 603 LANLVVVGSDRRKESKDLKEAKEMKMYELIETKJLNGQFRWISSQMRIRANVELYRYI 662
Qy 662 DTKGAFVQPALYEAFTGLTVEAMTCGPTFATCNGGPAEIIYHKGSGFNIDPYHGDOAA 721
Db 663 DTKGAFVQPALYEAFTGLTVEAMTCGPTFATCNGGPAEIIYHKGSGFNIDPYHGDOAA 722
Qy 722 ILVDFEKKCKDDPSHMWKISQGLKRIEKKYTWKIYSERLTLTGYYGFMKHYVSNLERR 781
Db 723 ILVDFEKKCKDDPSHMWKISQGLKRIEKKYTWKIYSERLTLTGYYGFMKHYVSNLERR 782
Qy 782 ESRRYLEMFPALYKRYKLAESVPLAE 806
Db 783 ESRRYLEMFPALYKRYKLAESVPLAE 807
```

RESULT 3

```
YUPOS
sucrose synthase (EC 2.4.1.13) - potato
C:Species: Solanum tuberosum (potato)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 11-Jun-1999
C:Accession: A29615
R:Salanoubat, M.; Belliard, G.
Gene 60, 47-56, 1987
A:Title: Molecular cloning and sequencing of sucrose synthase cDNA from potato (Solanum
A:Reference number: A29615; MUID:88152501; PMID:2964386
A:Accession: A29615
A:Molecule type: mRNA
A:Residues: 1-805 <SBL>
A:Cross-references: GB:M18745; NID:G169571; PIDN:AAA33841.1; PID:G169572
A:Experimental source: var. Sirtema 2n=4x
C:Superfamily: sucrose synthase; sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:279-753/Domain: sucrose-phosphate synthase homology <SSPS>
```

```
Query Match 82.4%; Score 3501.5; DB 1; Length 805;
Best Local Similarity 82.1%; Pred. No. 3.8e-224;
Matches 662; Conservative 65; Mismatches 78; Indels 1; Gaps 1;

Qy 1 MAERALLTVHSLRELRDELTLAHRNEILLASRIEKGKGILOHQIILEFEAIPENRKK 60
Db 1 MAERALLTVHSLRELRDELTLAHRNEILLASRIEKGKGILOHQIILEFEAIPENRKK 60
Qy 61 KLANGAFPEVLKASQDAIVLPWPVALAVRPFGWYIRVNVNVALVVEELVAEYLHFKEE 120
Db 61 KLANGAFPEVLKASQDAIVLPWPVALAVRPFGWYIRVNVNVALVVEELVAEYLHFKEE 120
Qy 121 ELVDGSSNGNFVLELDFEPNNSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMHPLE 180
Db 121 ELVDGSSNGNFVLELDFEPNNSFPRTLSKISGNGVEFLNRHLSAKLPHDKESMHPLE 180
```


A:Map position: 9
A:Introns: 32/2; 123/2; 187/3; 227/2; 299/3; 331/3; 389/3; 428/3; 484/2; 559/2; 66
C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:276-750/Domains: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 77.0%; Score 3268.5; DB 1; Length 802;
Best Local Similarity 76.2%; Pred. No. 1.1e-208;
Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;

OY 6 LTRVHSLEPERDITLAHNEILALSRIBEGKGIILOHOLIEFEALPEENKRLANG 65
DB 5 LTRHSLERLGRATSSHPNELIALFSRYVHGKMLQRHQLAEFDLFDSDKEKTA-- 62
OY 66 AFPEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVVEELTVAEYLFKEELVDG 125
DB 63 PFEDILRAQBAIVLPWVALAIRRPGVMDYIRNVSELAVELSEYIALAFKEQLVDG 122
OY 126 SSNGNFVLELDPEFPNSSPPRPPLTSKISGNGVEFLNRHLSAKLPDKESMHPLEPLRVH 185
DB 123 QSNSEFVLELDPEFPNASSFPSPMSKISGNGVQFLNRHLSKLFQDKESLVLPLNFKAH 182
OY 186 CHKGNNMLNRIQVNLALQHVLRKAERYLGLPPEPCAEFEHRFOEIGLERGMDTAE 245
DB 183 NYKGTVMNDRISQSLRGLQSLRAAEYLLSVPODTYSEFNHRFOELGKEKGMDTAK 242
OY 246 RVLNMIQLDLLEADPTCLEKFLGRIPMVFNVLTPHGFADNVLYGPDTCGGVVY 305
DB 243 RVLDTLHLLDLLEAPDANLEKFLGTPMVFNVLTPHGFADNVLYGPDTCGGVVY 302
OY 306 ILDOVRALNEMMLRIKQOGLNTPRLLITRLLPDAVGTTCQGRLEKYGTSHSDILRV 365
DB 303 ILDOVRALNEMMLRIKQOGLNTPRLLITRLLPDAVGTTCQGRLEKYGTSHSDILRV 362
OY 366 PPRTEKGIYVRKWIISFEKVPYLETTEYEDVAHEISKELHGPDLITGNXSGNIVASLLA 425
DB 363 PPRTEKGIYVRKWIISFEKVPYLETTEYEDVAHEISKELHGPDLITGNXSGNIVASLLA 421
OY 426 HKLGVTQCTIAHALEKTYKPPSDIYWKLEDDKYHPSQCFADLFAMNHTDFTITSTFOEI 485
DB 422 HKLGVTQCTIAHALEKTYKPPSDIYWKLEDDKYHPSQCFADLFAMNHTDFTITSTFOEI 481
OY 486 AGSKDTVQYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPTTEKRRLLKH 545
DB 482 AGSKDTVQYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPTTEKRRLLKH 541
OY 546 HPEIEDLLTYKVENNEHCYLANDRKRPILFTMPRLDRKONLGLVENKGPVKLRRLANL 605
DB 542 HPEIEELLYSDEVSEHNFVLKDKKRPILFTMPRLDRKONLGLVENKGPVKLRRLANL 601
OY 606 VVVGDRRESKDLSEKAKMKMFFELIDKYNLNGQFRWISSQMNIRVNLVRYICTDTCG 665
DB 602 VIVAGDHKESKDRBOAHFKKMYSLIDBYLKGIIIRISIAQMRVRRGELTRYICTDTCG 661
OY 666 AFVQALYEAAGLTVVEAMTCGLPTFAICNGGPAAIIVHKSGFNIDYHSDQAADILVD 725
DB 662 AFVQALYEAAGLTVVEAMTCGLPTFAICNGGPAAIIVHKSGFNIDYHSDQAADILVD 721
OY 726 FPEKKKDPSSHDKTSOGGLKRIEYKTYKTYISELLTLTGYYGWMKVSUNLERESRY 785
DB 722 FPDCKKADPSYWDIEISOGGLQRIYKTYKTYISELLTLTGYYGWMKVSUNLERESRY 781
OY 786 LEMFYALKYRKLAESVPLA 804
DB 782 LEMFYALKYRKLAESVPLA 800

RESULT 6
S19139
sucrose synthase (EC 2.4.1.13) 2 - rice
N:Alternate names: sucrose-UDP glucosyltransferase 2
C:Species: Oryza sativa (rice)
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jun-1999

C:Accession: S19139, S77947, S77943; S22536
R:Yu, W.P.; Wang, A.Y.; Huang, R.H.; Sung, H.Y.; Su, J.C.
Plant Mol. Biol. 18, 139-142, 1992
A:Title: Isolation and sequences of rice sucrose synthase cDNA and genomic DNA.
A:Reference number: S19139; MUID:92119223; PMID:1531032
A:Accession: S19139

A:Molecule type: DNA
A:Residues: 1-816 <YU>
A:Cross-references: EMBL:X59046
A:Accession: S77947
A:Molecule type: mRNA
A:Residues: 1-816 <YU>
A:Cross-references: EMBL:X59046
R:Su, J.C.
submitted to the EMBL Data Library, April 1991
A:Reference number: S77943
A:Accession: S77943

A:Molecule type: DNA
A:Residues: 1-447, 'Y', 449-816 <SU>
A:Cross-references: EMBL:X59046; NID:920094; PIDN:CAA41774.1; PID:920095
R:Wang, A.Y.; Yu, W.P.; Huang, R.H.; Sung, H.Y.; Su, J.C.
Plant Mol. Biol. 18, 139-142, 1992
A:Title: Presence of three rice sucrose synthase genes as revealed by cloning and sequen

A:Reference number: S22535; MUID:92288314; PMID:1534703
A:Accession: S22536
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 605-816 <WAN>
A:Experimental source: cv. Tainong 67
C:Genetics:
A:Introns: 37/2; 80/3; 131/2; 195/3; 235/2; 307/3; 339/3; 397/3; 436/3; 492/2; 567/2; 67
C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:284-758/Domains: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 76.9%; Score 3264; DB 2; Length 816;
Best Local Similarity 76.0%; Pred. No. 2.2e-208;
Matches 612; Conservative 91; Mismatches 100; Indels 2; Gaps 2;

OY 3 ERALTVRHSLEPERDITLAHNEILALSRIBEGKGIILOHOLIEFEALPEENKRL 61
DB 7 DRVLSRLHSVNERIGDSLSAHNPDLVAVFTLVNLGKMLQAHQIIAEYNNIAISEADKX 66
OY 62 LANGAFPEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVVEELTVAEYLFKEE 121
DB 67 LKQCAFEDVLKASQEAIVLPWVALAIRRPGVMEYIRVNVSELAVELTPEYLOFKEQ 126
OY 122 LVDOSSNGNFVLELDPEFPNSSPPRPPLTSKISGNGVEFLNRHLSAKLPDKESMHPLEPLRVH 181
DB 127 LVDOSSNGNFVLELDPEFPNASSFPSPMSKISGNGVQFLNRHLSKLFQDKESLVLPLNFKAH 186
OY 182 LRVHCHGKNMMLNDRIONLALQHVLRKAERYLGLPPEPCAEFEHRFOEIGLERGMDTA 241
DB 187 LRVHCHGKNMMLNDRIONLALQHVLRKAERYLGLPPEPCAEFEHRFOEIGLERGMDTA 246
OY 242 DTAERVLEMIQLDLLEADPTCLEKFLGRIPMVFNVLTPHGFADNVLYGPDTCGGVVY 301
DB 247 DTAERVLEMIQLDLLEADPTCLEKFLGRIPMVFNVLTPHGFADNVLYGPDTCGGVVY 306
OY 302 QVAVYILDOVRALNEMMLRIKQOGLNTPRLLITRLLPDAVGTTCQGRLEKYGTSHSDILRV 361
DB 307 QVAVYILDOVRALNEMMLRIKQOGLNTPRLLITRLLPDAVGTTCQGRLEKYGTSHSDILRV 366
OY 362 ILRVPPREKGIYVRKWIISFEKVPYLETTEYEDVAHEISKELHGPDLITGNXSGNIVASLLA 421
DB 367 ILRVPPREKGIYVRKWIISFEKVPYLETTEYEDVAHEISKELHGPDLITGNXSGNIVASLLA 425
OY 422 SLTAAKLGVTQCTIAHALEKTYKPPSDIYWKLEDDKYHPSQCFADLFAMNHTDFTITST 481
DB 426 SLTAAKLGVTQCTIAHALEKTYKPPSDIYWKLEDDKYHPSQCFADLFAMNHTDFTITST 485
OY 482 FOETIAGSKDTVQYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPTTEKRR 541

Db 486 FOEINAGNDTVGQYSESHNAFTWPGGLRVVHGIDVDPKFNIVSPGADMSIYEPYSESRRK 545
 Qy 542 LKHPRPETEDLLYTVEENEHCYANDRNKPLFTMPRLDRYKNTGLVENCGRPKARE 601
 Db 546 LTSHPETIELLYSSVDNNHFKMLKDRNKPIIFSMARLDRYKNTGLVELYGRNPLQOE 605
 Qy 602 LANLVVGGDRRKESKDLSEKAKMKMFELIDKYNLNGQFWMISSQMRIRINVELYRYIC 661
 Db 606 LANLVVGGDHGNPKDKEQAEQKPFKMFIDLEQYVNLNGHIRKISQMRIRINVELYRYIC 665
 Qy 662 DTGAFVOPALYEAAGLTVEEAMTGLPTFATCNGGPAEIIYHGKSGENIDPYHQDQAD 721
 Db 666 DTGAFVOPALYEAAGLTVEEAMTGLPTFATCNGGPAEIIYHGKSGENIDPYHQDQAD 725
 Qy 722 ILVDFEKKCKDPSHMDKISQGLKRIEKTWKYISERLLTLTGYYGFWKVSUMLERRE 781
 Db 726 LLVEFEKCKDPSHMTKISQGLKRIEKTWKYISERLLTLTGYYGFWKVSUMLERRE 785
 Qy 782 SRRYLFEMFYALKYRKLAESVPLAE 806
 Db 786 TRRYLEMLYALKYRKLAESVPLAE 810

RESULT 7

S23543
 sucrose synthase (EC 2.4.1.13) 1 - rice
 N:Alternate names: sucrose-UDP glucosyltransferase 1
 C:Species: Oryza sativa (rice)
 C:Date: 13-Jan-1999 #sequence_revision 13-Jan-1999 #text_change 18-Jun-1999
 C:Accession: S23543; #sequence_revision 13-Jan-1999 #text_change 18-Jun-1999
 R:Wang, M.B.; Boulter, D.; Gatehouse, J.A.
 Plant Mol. Biol. 19, 881-885, 1992
 A:Title: A complete sequence of the rice sucrose synthase-1 (RSS1) gene.
 A:Reference number: S23543; MUID:92153399; PMID:1386537
 A:Accession: S23543
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-808 <MAN>
 A:Cross-references: EMBL:X64770; NID:g20365; PIDN:CAA46017.1; PID:g20366
 R:Odgaard, W.; de Lumen, B.O.
 submitted to the EMBL Data Library, August 1992
 A:Description: Isolation and sequence of a sucrose synthase cDNA from developing rice
 A:Reference number: S25526
 A:Accession: S25526
 A:Molecule type: mRNA
 A:Residues: 1-190, 'P', 192-808 <ODE>
 A:Cross-references: EMBL:Z15028; NID:g20373; PIDN:CAA78747.1; PID:g20374
 C:Genetics:
 A:Introns: 32/2; 72/3; 123/2; 187/3; 227/2; 299/3; 331/3; 389/3; 428/3; 484/2; 559/2; 66
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:276-750/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 76.7%; Score 3259.5; DB 2; Length 808;
 Best Local Similarity 76.2%; Pred. No. 4.3e-208;
 Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;

Qy 6 LTRVHSLERLDELTLAARNELIALSRLEGKGILOHQIILFEFAIPEENRKLKANG 65
 Db 5 LARHSLERLDELTLAARNELIALSRLEGKGILOHQIILFEFAIPEENRKLKANG 62
 Qy 66 AFPEYLKASQEAIVLPVWALAVRRPGWMEYIRVNVHALVVEELTVAAEYLFKEELVDG 125
 Db 63 PFEDILRAQEAIVLPVWALAVRRPGWMEYIRVNVHALVVEELTVAAEYLFKEELVDG 122
 Qy 126 SSNGFVLELDEPEPNSFPPTLSKISGNGVEFLNRHLSAKLPHDKSMPLLEFLVH 185
 Db 123 HTNSFVLELDEPEPNSFPPTLSKISGNGVEFLNRHLSAKLPHDKSMPLLEFLVH 182
 Qy 186 CHKGKMMALNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFENRFOEIGLERGMDTA 245
 Db 183 NHKGTMMALNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFENRFOEIGLERGMDTA 242

Qy 246 RVLEMIQLDLLLEATDPTCELEKFLGRIPMVFNVLITPHGYAODNVLYGYPDTGGQVY 305
 Db 243 RVLDITLHLLDLLEAPDPALEKFLGRIPMVFNVLITPHGYAODNVLYGYPDTGGQVY 302
 Qy 306 ILDOVALENEMMLRIKQGINTPRILITRLLPDAVGTTCQGLREKVVGTSHSDILRV 365
 Db 303 ILDOVALENEMMLRIKQGINTPRILITRLLPDAVGTTCQGLREKVVGTSHSDILRV 362
 Qy 366 PPRFEKIVKRVISPFEEKVPYLETYEDVAHEISKELHGPDLITGXSGCNIVASLLA 425
 Db 363 PPRFENGLIRKVISPFEEKVPYLETYEDVAHEISKELHGPDLITGXSGCNIVASLLA 421
 Qy 426 HRLGVTOCTIAHALEKTYPPSDIYWKLEJRKHPSCOPTADLPAMNTDPIITSTFOEI 485
 Db 422 HRLGVTOCTIAHALEKTYPPSDIYWKLEJRKHPSCOPTADLPAMNTDPIITSTFOEI 481
 Qy 486 AGSKDTVQYSESHNAFTWPGGLRVVHGIDVDPKFNIVSPGADMSIYEPYSESRRK 545
 Db 482 AGSKDTVQYSESHNAFTWPGGLRVVHGIDVDPKFNIVSPGADMSIYEPYSESRRK 541
 Qy 546 HPEIEDLLYTVEENEHCYANDRNKPLFTMPRLDRYKNTGLVENCGRPKARELANL 605
 Db 542 HPEIEELLYSEVDEHFKVLKDRNKPIIFSMARLDRYKNTGLVENCGRPKARELANL 601
 Qy 606 VVVGDRRKESKDLSEKAKMKMFELIDKYNLNGQFWMISSQMRIRINVELYRYICDTKG 665
 Db 602 VVVGDRRKESKDLSEKAKMKMFELIDKYNLNGQFWMISSQMRIRINVELYRYICDTKG 661
 Qy 666 AFVOPALYEAAGLTVEEAMTGLPTFATCNGGPAEIIYHGKSGENIDPYHQDQADILVD 725
 Db 662 AFVOPALYEAAGLTVEEAMTGLPTFATCNGGPAEIIYHGKSGENIDPYHQDQADILVD 721
 Qy 726 FFEKCKDPSHMDKISQGLKRIEKTWKYISERLLTLTGYYGFWKVSUMLERRE 785
 Db 722 FFEKCKDPSHMDKISQGLKRIEKTWKYISERLLTLTGYYGFWKVSUMLERRE 781
 Qy 786 LEMFYALKYRKLAESVPLA 804
 Db 782 LEMFYALKYRKLAESVPLA 800

RESULT 8

T14338
 sucrose synthase (EC 2.4.1.13) isoform II - carrot
 C:Species: Daucus carota (carrot)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T14338
 R:Stum, A.; Leinhard, S.; Schatt, S.; Hardeggar, M.
 Plant Mol. Biol. 39, 349-360, 1999
 A:Title: Tissue-specific expression of two genes for sucrose synthase in carrot (Daucus
 A:Reference number: Z17990; MUID:99178785; PMID:10080700
 A:Accession: T14338
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-801 <STU>
 A:Cross-references: EMBL:Y16091; NID:g2760540; PIDN:CAA76057.1; PID:g2760541
 A:Experimental source: Cultivar Namtalse; mature leaves
 C:Genetics:
 A:Gene: Suzy*Dc2
 A:Introns: 29/2; 71/3; 122/2; 186/3; 298/3; 330/3; 388/3; 427/3; 483/2; 558/2; 746/2; 79
 C:Function:
 A:Description: catalyzes the reversible cleavage of sucrose into UDP-glucose and D-fructo-
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:275-749/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 76.3%; Score 3238.5; DB 2; Length 801;
 Best Local Similarity 76.1%; Pred. No. 1e-206;
 Matches 604; Conservative 85; Mismatches 104; Indels 1; Gaps 1;

Qy 11 SLRERLDELTLAARNELIALSRLEGKGILOHQIILFEFAIPEENRKLKANGAFPEV 70
 Db 7 SLRERLDELTLAARNELIALSRLEGKGILOHQIILFEFAIPEENRKLKANGAFPEV 66

[illegible][illegible]

A:Molecule type: mRNA
 A:Residues: 1-807 <HOZ>
 A:Cross-references: EMBL:X65871; NID:g19105; PIDN:CAA46701.1; PID:g19106
 C:Genetics:
 A:Map position: 7H
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:276-749/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 75.6%; Score 3209; DB 2; Length 807;
 Best Local Similarity 75.8%; Pred. No. 9.5e-205;
 Matches 606; Conservative 89; Mismatches 100; Indels 4; Gaps 3;

```

QY 6 LTRVSLERLDETLAAHNEITALLSRLEGKGLQHOHQIILFEPAIPENRK-KLANG 65
DB 5 LTRVSLERLDETLAAHNEITALLSRLEGKGLQHOHQIILFEPAIPENRK-KLANG 62
QY 66 AFEEVLKASQEAIVLPWVALAVRPRGWEYIRVNVHVALVEELTVAEYLHFKELVDG 125
DB 63 PFEDILRAAQAIAIVLPWVALAIRPTGVWDYIRVNVSELAVBELTVSLAFKQQLVDE 122
QY 126 SSNGFVLELDEFPNSSFPRPTLSKISGVNVEFLNRHLSAKLPHDKSMHPLLEFLRVH 185
DB 123 HASRFVLELDEFPNSSFPRPTLSKISGVNVEFLNRHLSAKLPHDKSMHPLLEFLRVH 182
QY 186 CHKGNMMLNDRIQNLNLQHLAKAEYIGTLPPETPCAEFEHRFOEIGLERGMDTA 245
DB 183 NYKGTMLNDRIQNLNLQHLAKAEYIGTLPPETPCAEFEHRFOEIGLERGMDTA 242
QY 246 RYLEMIQILLDLLEATDPTCTLEKFLGRIIPWVFNVLITPHGYFAQDNVLGYPTGGGVY 305
DB 243 RVHDITLHLLDLLEAPDPALEKFLGRIIPWVFNVLITPHGYFAQDNVLGYPTGGGVY 302
QY 306 ILDOVRALLENMLRIKQOGLNITPRILITRLLPDVAGTTGQRLKRYGTEHSDILRV 365
DB 303 ILDOVRALLENMLRIKQOGLNITPRILITRLLPDVAGTTGQRLKRYGTEHSDILRV 362
QY 366 PPRTEKGIYKRWKISFEKVPWYLETYEDVAHEISKELGTPDLITGNXSDNIVASILA 425
DB 363 PPRTEKGIYKRWKISFEKVPWYLETYEDVAHEISKELGTPDLITGNXSDNIVASILA 420
QY 426 HKLGVTQCTIAHALEKTKYPSDSIYWKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 485
DB 421 HKLGVTQCTIAHALEKTKYPSDSIYWKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 480
QY 486 AGSKDTVQYESHIAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPTYEERKRLKH 545
DB 481 AGSKDTVQYESHIAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPTYEERKRLKH 540
QY 546 HPETEDLLYTKVNEEHLCVLDRNKPILFTMPRLDRVKNLTGLVEMGCKNPKLRELANL 605
DB 541 HSEIELLYSDVENDEHKFVLKDRNKPILFTMPRLDRVKNLTGLVEMGCKNPKLRELANL 600
QY 606 VVVGDRRKESKDLBEKAKEMKMFELIDKYNLNGOFWMISQMRIRNVELYRYICDTKG 665
DB 601 VVVGDRRKESKDLBEKAKEMKMFELIDKYNLNGOFWMISQMRIRNVELYRYICDTKG 660
QY 666 AFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADILVD 725
DB 661 AFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADILVD 720
QY 726 FFEKCKKDP SHMDKISQGLKRIEKTWKIYSEBLTLTVGYGFWKIVSNLERRESRY 785
DB 721 FFEKCKKDP SHMDKISQGLKRIEKTWKIYSEBLTLTVGYGFWKIVSNLERRESRY 780
QY 786 LEMFALAKYRKLAESVPLA 804
DB 781 LEMFALAKYRKLAESVPLA 799

```

RESULT 11
 106497
 probable sucrose synthase (EC 2.4.1.13) 2 - garden pea

C:Species: Pisum sativum (garden pea)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
 C:Accession: T06497
 R:Buchner, P.
 submitted to the EMBL Data Library, October 1997
 A:Reference number: Z15720
 A:Accession: T06497
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-809 <BUC>
 A:Cross-references: EMBL:AJ001071; PIDN:CAA04512.1
 A:Experimental source: cultivar Filason
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:282-757/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 70.3%; Score 2985; DB 2; Length 809;
 Best Local Similarity 70.4%; Pred. No. 6.9e-190;
 Matches 565; Conservative 112; Mismatches 120; Indels 6; Gaps 4;

```

QY 7 TRVHSLERLDETLAAHNEITALLSRLEGKGLQHOHQIILFEPAIPENRK-KLAN 64
DB 8 TRVPSIRDRVDDTLAAHNEITALLSRVAAQGLQHOHQIILFEPAIPENRK-KLAN 67
QY 65 GAFFEVKASQEAIVLPWVALAVRPRGWEYIRVNVHVALVEELTVAEYLHFKELVDG 124
DB 68 GFPGIINSQAQIAIVLPWVALAIRPTGVWDYIRVNVSELAVBELTVSLAFKQQLVDE 127
QY 125 SSNGFVLELDEFPNSSFPRPTLSKISGVNVEFLNRHLSAKLPHDKSMHPLLEFLRVH 184
DB 128 GSKNDNITLLELDEFPNSSFPRPTLSKISGVNVEFLNRHLSAKLPHDKSMHPLLEFLRVH 187
QY 185 HCHKGNMMLNDRIQNLNLQHLAKAEYIGTLPPETPCAEFEHRFOEIGLERGMDTA 244
DB 188 HCHKGNMMLNDRIQNLNLQHLAKAEYIGTLPPETPCAEFEHRFOEIGLERGMDTA 247
QY 245 ERYLEMIQILLDLLEATDPTCTLEKFLGRIIPWVFNVLITPHGYFAQDNVLGYPTGGGVY 304
DB 248 ERYLEMIQILLDLLEATDPTCTLEKFLGRIIPWVFNVLITPHGYFAQDNVLGYPTGGGVY 307
QY 305 ILDOVRALLENMLRIKQOGLNITPRILITRLLPDVAGTTGQRLKRYGTEHSDILRV 364
DB 308 ILDOVRALLENMLRIKQOGLNITPRILITRLLPDVAGTTGQRLKRYGTEHSDILRV 367
QY 365 VPRTEKGIYKRWKISFEKVPWYLETYEDVAHEISKELGTPDLITGNXSDNIVASILA 424
DB 368 VPRTEKGIYKRWKISFEKVPWYLETYEDVAHEISKELGTPDLITGNXSDNIVASILA 426
QY 425 AHKLGVTQCTIAHALEKTKYPSDSIYWKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 484
DB 427 AHKLGVTQCTIAHALEKTKYPSDSIYWKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 486
QY 485 IAGSKDTVQYESHIAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPTYEERKRLKH 544
DB 487 IAGSKDTVQYESHIAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPTYEERKRLKH 546
QY 545 FHEPELIDLLYTKVNEEHLCVLDRNKPILFTMPRLDRVKNLTGLVEMGCKNPKLRELAN 604
DB 547 FHEPELIDLLYTKVNEEHLCVLDRNKPILFTMPRLDRVKNLTGLVEMGCKNPKLRELAN 606
QY 605 LVVVG--DRRKESKDLBEKAKEMKMFELIDKYNLNGOFWMISQMRIRNVELYRYICD 662
DB 607 LVVVG--DRRKESKDLBEKAKEMKMFELIDKYNLNGOFWMISQMRIRNVELYRYICD 665
QY 663 TKGAFFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADILVD 722
DB 666 TKGAFFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADILVD 725
QY 723 LVDFEKKCKDP SHMDKISQGLKRIEKTWKIYSEBLTLTVGYGFWKIVSNLERRES 782
DB 726 LVDFEKKCKDP SHMDKISQGLKRIEKTWKIYSEBLTLTVGYGFWKIVSNLERRES 785
QY 783 RRYLEMFALAKYRKLAESVPLA 805

```

[illegible]

Db	660	RYIADTRGAFAPAFYEAARGLTVLEAMTCGLPFTACHGDPABEIIIEHGLSGFHIDPYHPE	719
Qy	718	CAADILVDFEKKCKDPSHWDKISQGLKRIEIKTYKTIYSERLLTLGVYGFPMKHSNLL	777
Db	720	QAGNIMADPFEECKEDPNMKKVS DAGLQRIYERYTKIYSERIMTLAGVYGFPMKYVSKL	779
Qy	778	ERRRSRRYLEMRYALKYRKLAEISVP	802
Db	780	ERRRTRYLEMRYILKFRDLVKTVP	804
RESULT 13			
	571493	sucrose synthase (EC 2.4.1.13) - beet	
		CSpecies: Beta vulgaris (beet)	
		CDate: 09-Dec-1997 #sequence _revision 09-Dec-1997 #text_change 20-Jun-2000	
		CAccession: S71493; S71494	
		R.Hesse, H.; Willmitzer, L.	
		Plant Mol. Biol. 30, 863-872, 1996	
		ATitle: Expression analysis of a sucrose synthase gene from sugar beet (Beta vulgaris L.	
		AReference number: S71493; PMID:96270366; PMID:8639746	
		AAccession: S71493	
		A.Molecule type: mRNA	
		A.Residues: 1-822 <HES>	
		A.Cross-references: EMBL:X81974	
		R.Hesse, H.; Willmitzer, L.	
		submitted to the EMBL Data Library, September 1994	
		A.Description: Molecular cloning of a mitochondrial isoform of Cysteine Synthase from Ar	
		A.Reference number: S71494	
		AAccession: S71494	
		A.Molecule type: mRNA	
		A.Residues: 7'AG,-59-822 <HEM>	
		A.Cross-references: EMBL:X81974; NID:G1488569; PIDN:CAA57499.1; PID:G1488570	
		C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology	
		C.Keywords: glycosyltransferase; hexosyltransferase	
		F:280-755/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>	
Query Match	69.6%;	Score 2958;	DB 2; Length 822;
Best Local Similarity	70.3%;	Pred. No. 4.3e-168;	
Matches 566;	Conservative 106;	Mismatches 127;	Indels 6; Gaps 5;
Qy	6	LTRVNSLRRLDETLAHNEITALLSRIGSKGGLIOHQIILEFEALPRENRKK-LA	63
Db	5	LTRISMRRRVEDTSLVHNREIVSLSKYVAQGCLQPHNLIDGLSEVISBDKKQILUS	64
Qy	64	NGAFEFVLASQSEAIYLPFWVALAVRPRGVMEXIRNVNVALVVEBELVAEYLHFKEEYL	123
Db	65	DQPFSEVLRSQAEIAYVPPVALAVRPRGVMEXIRNVNVALVVEBELVAEYLHFKEEYL	124
Qy	124	DGSSNGNFVLELDFEEDFNSSFPRTLSKISGNGVEFLNRHLISAKLFHDKESMHPLEFLR	183
Db	125	DGKADHDVYLEDPEFPNESVPRAPTRSSSIGNGVQFNRHLSMFCNKDCLEPLDFLR	184
Qy	184	VHCHGKMMMLNDRIQNLNALQHLAKAEYLGTLPRETPAFEEHRRQEGLEGEGMDT	243
Db	185	VKHKGVMVMMNDRIQTLRIQLQSLSKAEEDYLLKLPADTPYSSEFPVLYQMGFEKGMGDT	244
Qy	244	ARRVLEMIDLLDLLEATPCTLEKFLGRIPMVFNVIILFPHYFAODNVLYGYPDTGGCV	303
Db	245	ARRVLEMMMLLDLIDAPDSTILETHIGRLPMVFNVIILSHGIFGQAHVGLPRTGGQI	304
Qy	304	VYILDQVRALENEMLRIRKQGGINLTPRLLITRLDPDVGTTGQRLKELYGTESDIL	363
Db	305	VYILDQVRLSHNMLORIKKGGLDVPRILIVSLRIDAGTTGNQRMKESVGTENASIL	364
Qy	364	RVPFRTKEGIIVKRWISRFEEKWVPLETYTDEDVAHSEKELHGTDLILGNXSDGNIYASL	423
Db	365	RVPFSEKEGILIRKWRISRPD-VWPLETFETEDAGGEIIGELQGRDILLIGNYSDGNIYASL	423
Qy	424	LAKHKGVTQCTTAHALEKTKYRPSDIIYKKLEDDKHSFCOPTALLFAMNHDFIITSTFQ	483
Db	424	LKHKGVTQCTTAHALEKTKYRPSDIIYKKRFEEDKHSFCOPTADMANNHDFIITSTIYQ	483

D06	147	SSS1GNGVQVJNVNRHLSISIMFRNKSESMEPLLEFLTRKHDGRPMMLNDR1QINP1ILOALA	206
QY	210	KAEYUJCTLPPEPPCMEFFHREFOEIGLEBGMGDTERLEWLTQILDLLEBTPCTJEKT	269
D06	207	RAEEFSLKPLAPLAPYSEFEELOQMGFEFGMDTQAKVSEVWHLLDLILOAPDPSVLETF	266
QY	270	LGRIPMFENVVILITPPGYFAODNVGLGPRYTGQVVYIILDOVALANEMTLIRIKOQGNIT	329
D06	267	LGRIPMFENVVILSRGYFPAQNVGLSPRTGQVVYIILDOVALANEMTLIRIKOQGLEVI	326
QY	330	PRILIIITRLLPDAVGTTCQRLKQYVTEHSDILVPRTEKGIYRWKMSIFSEKWPYLE	369
D06	327	PKLIIYTRLLPBAKGTTCNQRLERVSGTSHAHLIRPERTEKGIIRKXISRFD-VWPLYE	365
QY	390	TYTEDVAHEISKELHGTDRLLIGNSDGVIVASLAAHLCTYQCTIAALAEKTYPRSDI	449
D06	386	TFAPEDASNEISAELOQVPMILIGNSDGVIVASLAAHLCTYQCTIAALAEKTYPRSDI	445
QY	450	YMKLEEDKYFSCQFTADYFAMNHFTFITSTFOEIASKDTVQGYESHETFTLPGLYRV	509
D06	446	YMRNHEKHYFSSQFTADLMMNADFITSTYQETIASSKNVQGYESHETFTMPGLYRV	505
QY	510	VHGIVDFEPKFNVSFGADMEIYPPYTEEKRLKHPHEIEDLLYTKVENEHLVJNDR	569
D06	506	VHGIVDFEPKFNVSFGADMTIYFPYSDKERLTLHESIIEELLFSAQONDEHVGILLSDQ	565
QY	570	NKPLLFMPRLDVKVNLTGJVEKCGNPRLETLAVLVVVG-DRKSKXOLEEYAEKKM	628
D06	566	SKPLIFSMALDVKVNLTGJVECYANSLKRELAVLVVVGIDIBENOSRDEEAAELQKM	625
QY	629	FELIDKYNLNGQFRMISQMNRIARNVELRYVCDPKGAFVOPALYEAAGLVVYAMTCGL	688
D06	626	HSLLIEQDLDHGERFMTLAQONRNVRNDELRYIADTKGVFGVPAFPAEARGLVVYESMTCAL	685
QY	689	PTPATCNGSPAEIIVHGKSGFNIDPYHGOAADILVDFEKKCKDPSHWDKISOGGLKRI	748
D06	686	PTPATCNGSPAEIIEENGSGFNIDPYHPOVAGSLA-LFECNTNPMNHWKISISGGGLKRI	744
QY	749	EEKYTKWIVYEEBRLTLTGUYGFWKHNSUNERESRYYLMEFPALKYRKLASVYLAEE	806
D06	745	YERTYTKWYSEBRLTLTAGVAFEMFKHYSKLERRETRYLMEFSLKPRDLANSIDPLAD	802

Search completed: June 2, 2004, 14:53:14
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: June 2, 2004, 14:47:54 ; Search time 18 Seconds
(without alignments)
2311.586 Million cell updates/sec

Title: US-10-003-405-2
4247

Perfect score: 1 MAERALTRVSLRERLDEL.....EMFYALKRKAEVPLAE 806

Sequence: Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3763.5	88.6	805	1	SUSY SOYBN
2	3714.5	87.5	805	1	SUSY PHAUV
3	3669.5	86.4	806	1	SUSY VICFA
4	3648.5	85.9	805	1	SUSY MEDSA
5	3589.5	84.5	806	1	SUSY ARATH
6	3515.5	82.8	803	1	SUSY ALNGL
7	3501.5	82.4	805	1	SUSY SOLTU
8	3496.5	82.3	805	1	SUSY LYCES
9	3476.5	81.9	805	1	SUSY SOLTU
10	3400.5	80.1	808	1	SUS1 DAUCA
11	3301	77.7	820	1	SUS2 TULGE
12	3295	77.6	816	1	SUS2 MAIZE
13	3273	77.1	805	1	SUS1 TULGE
14	3273	77.1	816	1	SUS2 ORYSA
15	3268.5	77.0	802	1	SUS1 MAIZE
16	3262	76.8	816	1	SUS3 ORYSA
17	3252.5	76.6	808	1	SUS1 ORYSA
18	3238.5	76.3	801	1	SUS2 DAUCA
19	3222	75.9	816	1	SUS2 HORVU
20	3209	75.6	807	1	SUS1 HORVU
21	2985	70.3	809	1	SUS2 PEA
22	2865.5	67.5	805	1	SUS2 ARATH
23	2841	66.9	766	1	SUSY BETVU
24	661	15.6	218	1	SUSY SACOF
25	437	10.3	1059	1	SPS VICFA
26	436	10.3	1068	1	SPS MAIZE
27	431	10.1	1056	1	SPS SPICOL
28	421.5	9.9	1045	1	SPS BETVU
29	418.5	9.9	1081	1	SPS2 CRAPL
30	415.5	9.7	1053	1	SPS SOLTU
31	407	9.6	1057	1	SPS1 CITUN
32	394	9.3	1054	1	SPS1 CRAPL
33	387.5	9.1	1049	1	SPS ORYSA

34	156	3.7	377	1	YPJH BACSU
35	148	3.5	5430	1	MACF_HUMAN
36	147	3.5	5938	1	Q96PK2 homo sapien
37	142	3.3	358	1	YC07_KLEPN
38	135.5	3.2	406	1	YAS9_METUA
39	131	3.1	5327	1	MACF_MOUSE
40	129	3.0	390	1	YG07_METUA
41	128	3.0	5171	1	BPEA_HUMAN
42	127	3.0	380	1	CAPM_STRAU
43	124	2.9	359	1	YQGM_BACSU
44	124	2.9	461	1	GP13_YEAST
45	123	2.9	1042	1	TIRH_METUA

ALIGNMENTS

RESULT 1
ID SUSY SOYBN STANDARD; PRT; 805 AA.
AC P13708; Q22624; 01-JAN-1990 (Rel. 13, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP-glucosyltransferase)
DE (Nodulin-100).
GN SS.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hobbit; TISSUE=Root nodules;
RA Zhang X.-Q., Verma D.P.S., Patel S., Arredondo-Peter R., Miao G.-H.,
RA Kuismanen R., Klueae R.V., Chollet R.;
RT "Cloning of a full-length sucrose synthase cDNA from soybean (Glycine
RT max) root nodules."
RL (In) Plant Gene Register PGR97-173.
RN [2]
RP SEQUENCE OF 679-805 FROM N.A.
RC STRAIN=cv. Prize; TISSUE=Root nodules;
RX MEDLINE=88033030; PubMed=2889731;
RA Thummler F., Verma D.P.S.;
RT "Nodulin-100 of soybean is the subunit of sucrose synthase regulated
RT by the availability of free heme in nodules."
RL J. Biol. Chem. 262:14730-14736 (1987).
RN [3]
RP PHOSPHORYLATION.
RX MEDLINE=97379292; PubMed=9237614;
RA Zhang X.-Q., Chollet R.;
RT "Seryl-phosphorylation of soybean nodule sucrose synthase
RT (nodulin-100) by a Ca2+-dependent protein kinase."
RL FEBS Lett. 410:126-130 (1997).
CC - FUNCTION: Sucrose-leaving enzyme that provides UDP-glucose and
CC fructose for various metabolic pathways.
CC - CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC - SUBUNIT: Homotrimer.
CC - PM: Phosphorylated on serine residue(s).
CC - SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
CC sucrose synthase subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC EMBL: AF030231; AAC39323.1; -
CC PIR: A29484; A29484.

DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1; 1.
 DR Pfam; PF00862; Sucrose synth. 1.
 KM Transferrase; Glycosyltransferase; Nodulation; Phosphorylation.
 FT MOD_RES 11 11 PHOSPHORYLATION (BY CDPK) (POTENTIAL).
 FT CONFLICT 680 680 V -> L (IN REF. 2).
 FT CONFLICT 737 738 DK -> ET (IN REF. 2).
 FT CONFLICT 804 804 A -> V (IN REF. 2).
 SQ SEQUENCE 805 AA; 92243 MW; BBI4E5SEB8F5587B CRC64;

Query Match 88.6%; Score 3763.5; DB 1; Length 805;
 Best Local Similarity 88.0%; Pred. No. 8.6e-235;
 Matches 709; Conservative 49; Mismatches 47; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRRLDETLAHRNEITLALSRIGKKGILQHOHIIIEFAIPENRK 60
 DB 1 MATRLTRVHSRLRDLDTLANNRNEITLALSRIGKKGILQHOHIIIEFAIPENRK 60
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVRPRGVMEYIRVNVHALVVEELTVAEYLHFK 120
 DB 61 KLTDGAFGEVLRSQEAIVLPWVALAVRPRGVMEYIRVNVHALVVEELQPAEYLHFK 120
 QY 121 ELVDGSSNGNFVLELDEPPFNSSPPRPTLSKISNGVEFLNRHLSAKLPHDKESMHPLE 180
 DB 121 ELVDGSSNGNFVLELDEPPFNAPRPTLNKISNGVOFLNRHLSAKLPHDKESMHPLE 180
 QY 181 FLRVCHGKXMMNLNDRIQNLALOHVLRKAEVLTLPETPCAEFEHRFOELGLERGW 240
 DB 181 FLRLSVGKXKMLNDRIQNPALOHVLRKAEVLTGPETTPSEFEKFOELGLERGW 240
 QY 241 GDTAEVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPPGYFAQDNVLGYPTDG 300
 DB 241 GDNARVLESIQLLDLLEAPDPTCLEFLGRIPMVFNVLITSPGYFAQDNVLGYPTDG 300
 QY 301 GGVVYIILDQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 DB 301 GGVVYIILDQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 QY 361 DILRVPRTKGIYKWKISRFKWPVLETYTEDVAHREISKEHOTPDLIIGNSDGNIV 420
 DB 361 HILRVPRTKGIYKWKISRFKWPVLETYTEDVAHREISKEHOTPDLIIGNSDGNIV 419
 QY 421 ASLNAKLGVTQCTIAHALEKTKYPSDSIYWKLEBDKXHFSCQFTADLFPAMNHTDFILTS 480
 DB 420 ASLNAKLGVTQCTIAHALEKTKYPSDSIYWKLEBDKXHFSCQFTADLFPAMNHTDFILTS 479
 QY 481 TFOBIAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKKNIVSGADMEIYFYTEER 540
 DB 480 TFOBIAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKKNIVSGADMEIYFYTEER 539
 QY 541 RLKHHPREIEDLYTKVNEBEHLCVLBNRNKPIFTMPRLDVKVLGTGVEMCKGNPKLR 600
 DB 540 RLTSHPREIEDLYTKVNEBEHLCVLBNRNKPIFTMPRLDVKVLGTGVEMCKGNPKLR 599
 QY 601 ELANLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGOFRRMISQMRIRINVELYRYI 660
 DB 600 ELVNVLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGOFRRMISQMRIRINVELYRYI 659
 QY 661 CDTKCAFVOPALYEAFLTVFAMTCGLPTFATCNGPAEIIIVHKSQFNIDPYHGDOAA 720
 DB 660 CDTKCAFVOPALYEAFLTVFAMTCGLPTFATCNGPAEIIIVHKSQFNIDPYHGDOAA 719
 QY 721 DILVVPFECKCKDPBHMWDKISOGGLKRIEKKYTWKISEBRLLTLTGAVGFPMHVSNNLRR 780
 DB 720 DILVVPFECKCKDPBHMWDKISOGGLKRIEKKYTWKISEBRLLTLTGAVGFPMHVSNNLRR 779
 QY 781 ESRRYLEMFYALKYRKLAESVPLAE 806
 DB 780 ESRRYLEMFYALKYRKLAESVPLAE 805

RESULT 2

SUSY PHAAU
 ID SUSY PHAAU STANDARD; PRT; 805 AA.
 AC 001390;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
 GN S51.
 OS Phaseolus aureus (Mung bean) (Vigna radiata).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 CC NCBI TaxID=3916;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 517-537.
 RA Arai M., Mori H., Imaeki H.;
 RT "Expression of the gene for sucrose synthase during growth of mung
 bean seedlings.";
 RL Plant Cell Physiol. 33:503-506(1992).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10266; BAA01108.1; -;
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1; 1.
 DR Pfam; PF00862; Sucrose synth. 1.
 KM Transferrase; Glycosyltransferase.
 SQ SEQUENCE 805 AA; 92092 MW; A39683CFD54EFA4 CRC64;

Query Match 87.5%; Score 3714.5; DB 1; Length 805;
 Best Local Similarity 87.3%; Pred. No. 1.2e-221;
 Matches 704; Conservative 47; Mismatches 54; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRRLDETLAHRNEITLALSRIGKKGILQHOHIIIEFAIPENRK 60
 DB 1 MATRLTRVHSRLRDLDTLANNRNEITLALSRIGKKGILQHOHIIIEFAIPENRK 60
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVRPRGVMEYIRVNVHALVVEELTVAEYLHFK 120
 DB 61 KLTDGAFGEVLRSQEAIVLPWVALAVRPRGVMEYIRVNVHALVVEELQPAEYLHFK 120
 QY 121 ELVDGSSNGNFVLELDEPPFNSSPPRPTLSKISNGVEFLNRHLSAKLPHDKESMHPLE 180
 DB 121 ELVDGSSNGNFVLELDEPPFNAPRPTLNKISNGVOFLNRHLSAKLPHDKESMHPLE 180
 QY 181 FLRVCHGKXMMNLNDRIQNLALOHVLRKAEVLTLPETPCAEFEHRFOELGLERGW 240
 DB 181 FLRLSVGKXKMLNDRIQNPALOHVLRKAEVLTGPETTPSEFEKFOELGLERGW 240
 QY 241 GDTAEVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPPGYFAQDNVLGYPTDG 300
 DB 241 GDNARVLESIQLLDLLEAPDPTCLEFLGRIPMVFNVLITSPGYFAQDNVLGYPTDG 300
 QY 301 GGVVYIILDQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 DB 301 GGVVYIILDQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 QY 361 DILRVPRTKGIYKWKISRFKWPVLETYTEDVAHREISKEHOTPDLIIGNSDGNIV 420
 DB 361 HILRVPRTKGIYKWKISRFKWPVLETYTEDVAHREISKEHOTPDLIIGNSDGNIV 419

```

QY 421 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEDEKHFSCQFTADLFANNHTDFIITS 480
DB 420 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEDEKHFSCQFTADLFANNHTDFIITS 479
QY 481 TFOEIASGSDTVGOYESHTAFTLPGIYRVVHIGIDVDFPKFNIYVSGADMEIYFPYTEER 540
DB 480 TFOEIASGSDTVGOYESHTAFTLPGIYRVVHIGIDVDFPKFNIYVSGADMEIYFPYTEER 539
QY 541 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFTMPRLDVKNLTGIVEMCGNPKLR 600
DB 540 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFTMPRLDVKNLTGIVEMCGNPKLR 599
QY 601 ELANLVVVGGRRKESKOLEEKAEMKMFELIDKYNLNGQFWMSSQNNRIRNVELYRYI 660
DB 600 ELVNLVVAAGDRRKESKOLEEKAEMKMFELIDKYNLNGQFWMSSQNNRIRNVELYRYI 659
QY 661 CDTKCAFQOPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIYHKGSGFNIDPHYGDQA 720
DB 660 ADTKCAFQOPAYEAFGLTVVEAMTCGLPTPATCNGGPAEIIYHKGSGFNIDPHYGDQA 719
QY 721 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVSNLERR 780
DB 720 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVSNLERR 779
QY 781 ESRRLVEMFYALKYRKLAESVPLAE 806
DB 780 ESRRLVEMFYALKYRKLAESVPLAE 805

```

RESULT 3
SUSY VICFA STANDARD; PRT; 806 AA.

```

ID AC P31926;
DT 01-JUL-1993 (Rel. 26; Created)
DT 01-JUL-1993 (Rel. 26; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
GN SUCS.
OS Vicia faba (Broad bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxId=3906;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Fribo; PubMed=7764025;
RX MEDLINE=93379571;
RA Heim U., Weber H., Baumlein H., Wobus U.;
RT "A sucrose-synthase gene of Vicia faba L.: expression pattern in
RT developing seeds in relation to starch synthesis and metabolic
RT regulation.";
RL Planta 191:394-401 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kleine Thueringer; TISSUE=Root nodules;
RX MEDLINE=94003420; PubMed=8400379;
RA Kuefer H., Fruhling M., Perlick A.M., Puchler A.;
RT "The sucrose synthase gene is predominantly expressed in the root
RT nodule tissue of Vicia faba.";
RL Mol. Plant Microbe Interact. 6:507-514 (1993).
CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
CC fructose for various metabolic pathways.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
CC sucrose synthase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; X69773; CAA9428.1; -.
DR EMBL; M97551; AAC37346.1; -.
DR PIR; S31479; S31479.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR000368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose synth; 1.
DR Transferrase; Glycosyltransferase.
SQ SEQUENCE 806 AA; 92520 MM; A51B004C0732F306 CRC64;

```

Query Match 86.4%; Score 3669.5; DB 1; Length 806;
Best Local Similarity 86.2%; Pred. No. 9.8e-229;
Matches 695; Conservative 52; Mismatches 58; Indels 1; Gaps 1;

```

QY 1 MAERALTVHSGRELDLTLAHRNEILLASRIEKGKGILOHQIILFEPAIPEERK 60
DB 1 MATERTLTVHSLRELDLTLAHRNEILLASRIEKGKGILOHQIILFEPAIPEERK 60
QY 61 KLANGAFPEVLKASGEATVLPWVALAVRPRGWMEYIRVNVHALLVVELTVAEYLHFK 120
DB 61 KLTDGAFGEVLKSTGEATVLPWVALAVRPRGWMEYIRVNVHALLVVELTVAEYLHFK 120
QY 121 ELVDGSSNGNFVLELDFEPFNSSFPPTLSKISNGVEFLNRHLSAKLFHDKESNHPLLE 180
DB 121 ELVDGSANGNFVLELDFEPFTASFPPTLSKISNGVGFNRHLSAKLFHDKESLHPLLE 180
QY 181 FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYIGTLPEPTPCAEFPHRQETGLENG 240
DB 181 FLRLHSYKGTMLMNDRIQNPDSLOHVLARKABEYISTVDPPEPPEPHRQETGLENG 240
QY 241 GDTARVLEMIQLLDLEATDPTCEKFLGRIPVNFVNLTPHVGPAQDNVGLGPTG 300
DB 241 GDSARVLESIQLLDLLEADPTCEKFLGRIPVNFVNLTPHVGPAQDNVGLGPTG 300
QY 301 GQVAVYILDQVBALENEMLRIRIKQOGLNITPRLITRLLPDAVGTTCQRIEKKYGTENS 360
DB 301 GQVAVYILDQVBALENEMLRIRIKQOGLDIPRILITRLLPDAVGTTCQRIEKKYGTENS 360
QY 361 DILRVPEPTEKGIYRKWISREKVPYLETEDVAHEISKELHSTPDLIIKNSDGNIV 420
DB 361 HILRVPEPDKGIYRKWISREKVPYLETEDVAHEISKELHSTPDLIIKNSDGNIV 419
QY 421 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEDEKHFSCQFTADLFANNHTDFIITS 480
DB 420 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEDEKHFSCQFTADLFANNHTDFIITS 479
QY 481 TFOEIASGSDTVGOYESHTAFTLPGIYRVVHIGIDVDFPKFNIYVSGADMEIYFPYTEER 540
DB 480 TFOEIASGSDTVGOYESHTAFTLPGIYRVVHIGIDVDFPKFNIYVSGADMEIYFPYTEER 539
QY 541 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFTMPRLDVKNLTGIVEMCGNPKLR 600
DB 540 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFTMPRLDVKNLTGIVEMCGNPKLR 599
QY 601 ELANLVVVGGRRKESKOLEEKAEMKMFELIDKYNLNGQFWMSSQNNRIRNVELYRYI 660
DB 600 ELVNLVVAAGDRRKESKOLEEKAEMKMFELIDKYNLNGQFWMSSQNNRIRNVELYRYI 659
QY 661 CDTKCAFQOPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIYHKGSGFNIDPHYGDQA 720
DB 660 CDTKCAFQOPAYEAFGLTVVEAMTCGLPTPATCNGGPAEIIYHKGSGFNIDPHYGDQA 719
QY 721 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVSNLERR 780
DB 720 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVSNLERR 779
QY 781 ESRRLVEMFYALKYRKLAESVPLAE 806
DB 780 ESRRLVEMFYALKYRKLAESVPLAE 805

```

RESULT 4

RA Rameberger U., Wedler H., Balke K., Wedler B., Peters S.,
 RA van Straten M., Dickse W., Moolman P., Klein Lankhorst R.,
 RA Weltzberger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lambrecht S., Villarreal R., Giesen J., Ardiles W.,
 RA Bents O., Lemke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
 RA Scheller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant *Arabidopsis*
 thaliana".
 RL Nature 408:823-826(2000).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X70990; CAA50317.1; -
 DR EMBL: AF296832; -; NOT ANNOTATED CDS.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glyco_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth; 1.
 DR Transferase: Glycosyltransferase; Multigene family.
 KM TRANSFERASE: Glycosyltransferase; Multigene family.
 FT CONFLICT 61 61 R -> Q (IN REF. 1).
 FT CONFLICT 108 108 V -> L (IN REF. 1).
 FT CONFLICT 222 222 S -> P (IN REF. 1).
 FT CONFLICT 405 405 N -> D (IN REF. 1).
 FT CONFLICT 434 434 C -> Q (IN REF. 1).
 FT CONFLICT 609 609 V -> I (IN REF. 1).
 FT CONFLICT 751 751 EK -> DE (IN REF. 1).
 FT CONFLICT 751 751 EK -> DE (IN REF. 1).
 SQ SEQUENCE 806 AA; 92798 MW; 201D365720D46A3F CRC64;
 Query Match 84.5%; Score 3569.5; DB 1; Length 806;
 Best Local Similarity 83.9%; Pred. No. 1.4e-223;
 Matches 675; Conservative 66; Mismatches 61; Indels 3; Gaps 2;
 QY 2 AERATRHVSREPERDETLAHRNEIALSRIBCKGKIIQHNOIIEFAIPENRKK 61
 DB 4 AERMTTRHVSQRENETLVSRNEVLAISLVEAKGKIIQOONQIIEFAIPQTRKK 63
 QY 62 LANGAFVEYLKASQSAIVLPWVALAVRPRGWMEXIRVNVVALVVEELTVAEYHFKEE 121
 DB 64 LEGGFPLDKSTQERIVLPWVALAVRPRGWMEXIRVNVVALVVEELTVAEYHFKEE 123
 QY 122 LVDGSSNGNFVLELDFEPNSSPPRPTLSKISGVNVEFLNRHLAKLPHDKESMPLLEF 181
 DB 124 LVDGVKNGNFLELDFEPFNASIPRPTLHKYIGNGVDFNRHLAKLPHDKESLPLLEF 183
 QY 182 LRVHKGKGMMLNDRIQNLALQVLRKAEVYLGTLPEPTCAEFERFQSIGERMG 241
 DB 184 LRLSHQGNLMLSEKIQNLNLQTLRLRAEYVLAELKSETLYEEFAEFIEIGERMG 243
 QY 242 DRAEVLVLEMIQILLDLLEATDPTLEKFLGRIPWVNVVILPHGVPADNVLAGYDPTGG 301
 DB 244 DNAEVLDVIRLLDLLEAPDCTLETFGRVPMVNVVILPHGVPADNVLAGYDPTGG 303
 QY 302 QVVVILIDQVRALENEMLRIRKQGGNITERIILIRLLPDAVGTTCGRLKRYGTEHSD 361
 DB 304 QVVVILIDQVRALEIEMLRIRKQGGNITERIILIRLLPDAVGTTCGERLERVYSEYCD 363
 QY 362 ILRVPRTERKGIIVRKWISFEKVPWPLETYTDDVAHSEIKELAGPDLILGNKSGNIVA 421
 DB 364 ILRVPRTERKGIIVRKWISFEKVPWPLETYTDDVAHSEIKELAGPDLILGNKSGNIVA 422
 QY 422 SLAAKLGVTOCTIAHALEKTKYPPDSIDYWKLEDEKHSQCFPTADLFAMNHTDPIITST 481

DB 423 SLAAKLGVTOCTIAHALEKTKYPPDSIDYWKLEDKYHFSQCFPTADLFAMNHTDPIITST 482
 QY 482 FOETAGSKDTGVGVESHAFITLPGIYRVVHGIDVDFDPENIVSPADMEIYFPYEEKR 541
 DB 483 FOETAGSKDTGVGVESHAFITLPGIYRVVHGIDVDFDPENIVSPADMEIYFPYEEKR 542
 QY 542 LKHFEPIEDLLTYTVENBEHLVYLDNRKILFTMPRLDRVKULTGLEWCGKPKLRE 601
 DB 543 LTRHSEIEBILLYSVENKEHLVYLDNRKILFTMPRLDRVKULTGLEWCGKPKLRE 602
 QY 602 LANLVVGGDRRKESKDLLEKAEKMKMELIDKYNLNGQFWMISQNMRRINVELRYIC 661
 DB 603 LANLVVGGDRRKESKDLLEKAEKMKMELIDKYNLNGQFWMISQNMRRINVELRYIC 662
 QY 662 DTGKAFVOPALYEAAGLTIVVEMTCGLPTPATCNGSPREIIVHSGSGNIDPHDQAD 721
 DB 663 DTGKAFVOPALYEAAGLTIVVEMTCGLPTPATCNGSPREIIVHSGSGNIDPHDQAD 722
 QY 722 ILVDFEKKDPSHMDKISQGLKRIEKTWKYISRLTLTGVSFWMKVSULERE 781
 DB 723 TLADFTCKEDPSHMDKISQGLKRIEKTWKYISRLTLTGVSFWMKVSULERE 782
 QY 782 SRRYLEMFYALKYRKLAESVPLAE 806
 DB 783 ARRYLEMFYALKYRKLAESVPLAE 805
 RESULT 6
 SUSY_ALNGL STANDARD; PRT; 803 AA.
 ID SUSY_ALNGL
 AC P49034;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
 GN SUS1.
 OS *Alnus glutinosa* (Alder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Euxrosida I; Fagales; Betulaceae; *Alnus*.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root;
 RX MEDLINE=96186875; PubMed=8602161;
 RA van Ghelue M., Pawlowski K.,
 RA Bisseling T., Paulsen H.,
 RT "Sucrose synthase: comparison with legume nodules".
 RT *Alnus glutinosa*: comparison with legume nodules".
 RL Mol. Gen. Genet. 250:437-446(1996).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92378; CAA63122.1; -
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glyco_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth; 1.
 DR Transferase: Glycosyltransferase.
 SQ SEQUENCE 803 AA; 91630 MW; 28C23736DDF1C731 CRC64;
 Query Match 82.8%; Score 3515.5; DB 1; Length 803;

Best Local Similarity 83.8%; Pred. No. 8,1e-219;
Matches 676; Conservative 56; Mismatches 70; Indels 5; Gaps 5;

OY 1 MAERALTVSHSLRELDDETLAHRNEITALLSRLEGKGILOHIOILFEALPEERK 60
DB 1 MAERLVTHSHSLRELDDETLVANRNEIYALLSRIGKGIICRNIOILAEVEALPEARX 60
OY 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
OY 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
OY 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
OY 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
OY 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
DB 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
OY 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
DB 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
OY 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300
DB 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300
OY 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300
DB 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300
OY 301 GQVVTILQVRLNEMILRIKQGLINTTPRLITLTPDAVGTTCQRLKRYGTSHS 360
DB 301 GQVVTILQVRLNEMILRIKQGLINTTPRLITLTPDAVGTTCQRLKRYGTSHS 360
OY 301 GQVVTILQVRLNEMILRIKQGLINTTPRLITLTPDAVGTTCQRLKRYGTSHS 360
DB 301 GQVVTILQVRLNEMILRIKQGLINTTPRLITLTPDAVGTTCQRLKRYGTSHS 360
OY 359 DILRVPRTEKIVQWISRRF-VMPRETYTEDVGVLEIKELQSKPDLIIGNSDGNIV 417
DB 359 DILRVPRTEKIVQWISRRF-VMPRETYTEDVGVLEIKELQSKPDLIIGNSDGNIV 417
OY 421 ASLALHKLGVTOCTAHALKEKTPDSDIYWKQKEDKXHFSCQTAADLPAAMHTDFIITS 480
DB 421 ASLALHKLGVTOCTAHALKEKTPDSDIYWKQKEDKXHFSCQTAADLPAAMHTDFIITS 480
OY 481 TFOEIASGKDTYGOESTHTAFPLBLYVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 540
DB 481 TFOEIASGKDTYGOESTHTAFPLBLYVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 540
OY 478 TFOEIASGKDTYGOESTHTAFPLBLYVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 536
DB 478 TFOEIASGKDTYGOESTHTAFPLBLYVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 536
OY 541 RLKHHPEIEDLLYKVENEBELCYLNDNRKPIELTMRPLDVKVLTGLVEMCGKNPKLR 600
DB 541 RLKHHPEIEDLLYKVENEBELCYLNDNRKPIELTMRPLDVKVLTGLVEMCGKNPKLR 600
OY 537 RLTSHPREIBELLYPVENEHELCTYKDRNRPITTMARLDVRKVIITGLVEMGKNTLR 596
DB 537 RLTSHPREIBELLYPVENEHELCTYKDRNRPITTMARLDVRKVIITGLVEMGKNTLR 596
OY 601 ELANLVVGGDRRKESKDLBEKAKEMKMFELIDKYNLNGOFMISSQMRIRANVELYRI 660
DB 601 ELANLVVGGDRRKESKDLBEKAKEMKMFELIDKYNLNGOFMISSQMRIRANVELYRI 660
OY 597 ELVNLVVAAGNLEKESKONBEKAMTKHGLIETKLNQOFMISSQMRIRANVELYRI 656
DB 597 ELVNLVVAAGNLEKESKONBEKAMTKHGLIETKLNQOFMISSQMRIRANVELYRI 656
OY 661 CDTKAP-VQPALYFAFGITVVEAMTCGPTPATCNGGPAEITVYKSGFNIDPHGQA 719
DB 661 CDTKAP-VQPALYFAFGITVVEAMTCGPTPATCNGGPAEITVYKSGFNIDPHGQA 719
OY 657 ADTKGGLCAGPAIYAFGLTVVESMTGCLPTFATCKGPAEITVYKSGFNIDPHGQA 716
DB 657 ADTKGGLCAGPAIYAFGLTVVESMTGCLPTFATCKGPAEITVYKSGFNIDPHGQA 716
OY 720 ADIIVDFEKKCKDSDHMDKISQGLKRIEKEYTWKISERLLTGLVYGFVKHSNLER 779
DB 720 ADIIVDFEKKCKDSDHMDKISQGLKRIEKEYTWKISERLLTGLVYGFVKHSNLER 779
OY 717 AOLVDFEKKCKDSDHMDKISQGLKRIEKEYTWKISERLLTGLVYGFVKHSNLER 776
DB 717 AOLVDFEKKCKDSDHMDKISQGLKRIEKEYTWKISERLLTGLVYGFVKHSNLER 776
OY 780 RESRRYLEMVALKYRKLAESVPLAE 806
DB 780 RESRRYLEMVALKYRKLAESVPLAE 806
OY 777 LESRRYLEMVALKYRKLAESVPLAE 803
DB 777 LESRRYLEMVALKYRKLAESVPLAE 803

RESULT 7
SUS1 SOLTU STANDARD; PRT; 805 AA.
ID SUS1 SOLTU STANDARD; PRT; 805 AA.
AC P10691;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase)
DE (SS16).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4113;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Sirema;
RX MEDLINE=88152501; PubMed=2964386;
RA Salanoubat M., Belliard G.;
RT "Molecular cloning and sequencing of sucrose synthase cDNA from
RT potato (Solanum tuberosum L.): preliminary characterization of
RT sucrose synthase mRNA distribution";
RL Gene 60:47-56 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. FL1607; TISSUE=leaf;
RX MEDLINE=96172787; PubMed=589622;
RA Fu H., Park W.D.;
RT "Sink- and vascular-associated sucrose synthase functions are encoded
RT by different gene classes in potato.";
RL Plant Cell 7:1369-1385 (1995).
CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
CC fructose for various metabolic pathways.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -1- TISSUE SPECIFICITY: Expression is at least 10 fold higher in
CC tubers compared to photosynthetically active tissues.
CC -1- SIMILARITY: Belongs to the glucosyltransferase family 1. Plant
CC sucrose synthase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.1db-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M18745; AAA33841.1; -;
DR EMBL: U24087; AAA97571.1; -;
DR PIR: A29615; YUFOS.
DR InterPro: IPR001296; Glyco_trans_1.
DR InterPro: IPR001368; Sucrose_synth.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Pfam: PF00862; Sucrose_synth_1.
CC Transferrase, Glycosyltransferase; Multigene family.
FT CONFLICT 221 221 E -> D (IN REF. 2).
FT CONFLICT 731 732 K D -> R (IN REF. 2).
FT CONFLICT 741 741 M -> T (IN REF. 2).
FT CONFLICT 748 748 E -> Q (IN REF. 2).
FT CONFLICT 759 759 S -> R (IN REF. 2).
SQ SEQUENCE 805 AA; 92416 MW; C453363A7ACD2809 CRC64;

Query Match 82.4%; Score 3501.5; DB 1; Length 805;
Best Local Similarity 82.1%; Pred. No. 6.5e-218;
Matches 662; Conservative 65; Mismatches 78; Indels 1; Gaps 1;

OY 1 MAERALTVSHSLRELDDETLAHRNEITALLSRLEGKGILOHIOILFEALPEERK 60
DB 1 MAERLVTHSHSLRELDDETLVANRNEIYALLSRIGKGIICRNIOILAEVEALPEARX 60
OY 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
OY 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVKAQGEAIVLPPWVALAVRPPGWEXIRVVNVALVVEELTVAEYLHKE 120
OY 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
OY 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNSSFPRLTSSKISNGVEFLNRHLSAKLPHKESMHPLE 180
OY 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
DB 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
OY 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
DB 181 FLRVCHGKMMMLNDRIQNLALOHVLRKAEYVLTPTPEPCAEFERFOEIGLERGW 240
OY 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300
DB 241 GDTAERVLEMIQLDLLEATDPTCLEKFLGRIPMVFNVLITPHGYAODNVLGYPDTG 300

```

QY 301 GQVYVILDOVALRENMELRLIKOQGLNTTPRLITRLPRAVGTTCGOREKXYGTSHS 360
DB 301 GQVYVILDOVALRENMELRLIKOQGLNTTPRLITRLPRAVGTTCGOREKXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPLYETETEDVAHEISXELGTDPDIIIGNSDGNIV 420
DB 361 HILRVPTTEKGIYRKWISREPKWPLYETETEDVAHEISXELGTDPDIIIGNSDGNIV 420
QY 421 ASLAAHKLGVTQCTTAHAELEKTKYPSDSIYWKKEDEKXHFSCQPTADLIANNHTDFIITS 480
DB 421 ASLAAHKLGVTQCTTAHAELEKTKYPSDSIYWKKEDEKXHFSCQPTADLIANNHTDFIITS 480
QY 481 TFOEIASGKDVGVGYESHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEERK 540
DB 481 TFOEIASGKDVGVGYESHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEERK 540
QY 541 RLKHFPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
DB 541 RLKHFPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 601 ELANLVVVGDRRKESKOLEEAKEMKMFELIDKYNLNGQFRWISSQNNRIRANVELYRYI 660
DB 601 ELANLVVVGDRRKESKOLEEAKEMKMFELIDKYNLNGQFRWISSQNNRIRANVELYRYI 660
QY 661 CDTGAFVQPALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGKSGFNIDPYHGDOAA 720
DB 661 CDTGAFVQPALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTGLVYGFVKHVSNTERR 780
DB 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTGLVYGFVKHVSNTERR 780
QY 781 ESRRLMEFYALKYRKLAESVPLAE 806
DB 781 ESRRLMEFYALKYRKLAESVPLAE 806
QY 780 EIRRYLEMFYALKYRKMAEAVPLAE 805
DB 780 EIRRYLEMFYALKYRKMAEAVPLAE 805

```

RESULT 8

SUS2_LYCES STANDARD; PRT; 805 AA.

```

AC P490337;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP-glucosyltransferase).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VF36; TISSUE=Placit1;
RX MEDLINE=94120019; PubMed=8290642;
RA Wang F., Smith A.G., Brenner M.L.;
RT "Isolation and sequencing of tomato fruit sucrose synthase cDNA.";
RL Plant Physiol. 103:1463-1464 (1993).
CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
fructose for various metabolic pathways.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
sucrose synthase subfamily.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC EMBL; L19762; AAA34196.1; -
 DR InterPro; IPR001296; Glyco_trans_1.

DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucrose_synth_1.
 KM Transferase; Glycosyltransferase.
 SQ SEQUENCE 805 AA; 9246 MW; 8C7CCB094154835 CRC64;

Query Match 82.3%; Score 3496.5; DB 1; Length 805;
 Best Local Similarity 81.9%; Pred. No. 1.4e-217;
 Matches 660; Conservative 67; Mismatches 78; Indels 1; Gaps 1;

```

QY 1 MAERALRVHSLRRLDETLAARNEILLASRIEGKGGILOHQQIILEFALPEENRK 60
DB 1 MAERALRVHSLRRLDETLAARNEILLASRIEGKGGILOHQQIILEFALPEENRK 60
QY 61 KLANGAFPEVKAQOEAIVLPPWALAVRPFGWYIRRVNVALVVEELSVPELOPKE 120
DB 61 KLANGAFPEVKAQOEAIVLPPWALAVRPFGWYIRRVNVALVVEELSVPELOPKE 120
QY 121 ELYDSSNGNFVLEIDEPFSSFPRLTSSKISNGVGFNRHLSAKLFDHKSHPLE 180
DB 121 ELYDSSNGNFVLEIDEPFSSFPRLTSSKISNGVGFNRHLSAKLFDHKSHPLE 180
QY 181 FLRVYCHGKNMMDNRIONALQHVLRKAEVYGLTLPETPCAEEHREPOEIGLEKGM 240
DB 181 FLRVYCHGKNMMDNRIONALQHVLRKAEVYGLTLPETPCAEEHREPOEIGLEKGM 240
QY 241 GDTREVLTEMQLLDLEATDPCLEKFLGRIMVENVVLTTHGTFADQNVTCYPTG 300
DB 241 GDTREVLTEMQLLDLEATDPCLEKFLGRIMVENVVLTTHGTFADQNVTCYPTG 300
QY 301 GQVYVILDOVALRENMELRLIKOQGLNTTPRLITRLPRAVGTTCGOREKXYGTSHS 360
DB 301 GQVYVILDOVALRENMELRLIKOQGLNTTPRLITRLPRAVGTTCGOREKXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPLYETETEDVAHEISXELGTDPDIIIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREPKWPLYETETEDVAHEISXELGTDPDIIIGNSDGNIV 420
QY 421 ASLAAHKLGVTQCTTAHAELEKTKYPSDSIYWKKEDEKXHFSCQPTADLIANNHTDFIITS 480
DB 421 ASLAAHKLGVTQCTTAHAELEKTKYPSDSIYWKKEDEKXHFSCQPTADLIANNHTDFIITS 480
QY 481 TFOEIASGKDVGVGYESHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEERK 540
DB 481 TFOEIASGKDVGVGYESHTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEERK 540
QY 541 RLKHFPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
DB 541 RLKHFPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 600 GLVNLVVVGDRRKESKOLEEAKEMKMFELIDKYNLNGQFRWISSQNNRIRANVELYRYI 659
DB 600 GLVNLVVVGDRRKESKOLEEAKEMKMFELIDKYNLNGQFRWISSQNNRIRANVELYRYI 659
QY 661 CDTGAFVQPALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGKSGFNIDPYHGDOAA 720
DB 661 CDTGAFVQPALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTGLVYGFVKHVSNTERR 780
DB 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTGLVYGFVKHVSNTERR 780
QY 780 EIRRYLEMFYALKYRKMAEAVPLAE 805
DB 780 EIRRYLEMFYALKYRKMAEAVPLAE 805

```

RESULT 9

SUS2_SOLITU STANDARD; PRT; 805 AA.

AC P490337;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update).

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase)
 DE (S865).
 OS Solanum tuberosum (potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. FL1607; TISSUE=leaf;
 RX MEDLINE=96172787; PubMed=8589622;
 RA Fu H., Park W.D.;
 RT "Sink- and vascular-associated sucrose synthase functions are encoded
 by different gene classes in potato.";
 RL Plant Cell 7:1369-1385(1995).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glucosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U24088; AAA97572.1; -
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1.
 DR Pfam; PF00862; Sucrose_synth_1.
 DR Transferase; Glycosyltransferase; Multigene family.
 KW SEQUENCE 805 AA; 92577 MW; 6117DBD2E4DD35DR CRC64;
 SQ
 Query Match 81.9%; Score 3476.5; DB 1; Length 805;
 Best local similarity 81.9%; Pred. No. 2,7e-216;
 Matches 660; Conservative 65; Mismatches 80; Indels 1; Gaps 1;
 QY 1 MAERALTRVHSLSRERLDDELTAHNRNEIALLSRIGSKGKILQHNIIFEEALPEENRK 60
 DB 1 MAERLVTVHSLRERLDDELTAHNRNEIIFLSRISHGKILKPHQLAEFESIKHKD 60
 QY 61 KLANAFEEVYLKASGEATVLPWVALAVRPRGWVEYIRVNVHVALVVELTVAELVHKE 120
 DB 61 KLNDHAFEEVYKSTGEATVLPWVALAIRPGVVEYVRVNVNALIVEELTVPELQPE 120
 QY 121 ELVDSSNGNFVLELDPEFPNNSPPRLTSSKISGNGVEFLNRHLGAKLFHDKESMPLLE 180
 DB 121 ELVNGTSDNDFVLELDPEFPNNSPPRLTSSKISGNGVEFLNRHLGAKLFHDKESMPLLE 180
 QY 181 FLRVHCHGKQKMMNDRIQNLALQHLVLRKAEVYLGILPETPCAEEFHRFOEIGLERGW 240
 DB 181 FLRVHCHGKQKMMNDRIQNLALQHLVLRKAEVYLTITSPETSYSAPFHKKOEIGLERGW 240
 QY 241 GPTARVVEMLTOLLDLLEATPCTLEKFLGIRPVWVNVVITTPRGYADQNVLGYPRTG 300
 DB 241 GPTARVVEMLTOLLDLLEATPCTLEKFLGIRPVWVNVVITTPRGYADQNVLGYPRTG 300
 QY 301 GOVVVILDOVRLAENEMLRIRKOGANTPRLLITRLLPDAVGTTCORLEKVGSTHS 360
 DB 301 GOVVVILDOVRLAENEMLRIRKOGANTPRLLITRLLPDAVGTTCORLEKVGSTHS 360
 QY 361 DILRVPEFTEKGIYKWSIRFEKWPVYLETYTEVAHEISKELHGTPLLIGNSDGNIV 420
 DB 361 HILRVPEFTEKGIYKWSIRFE-VWPMYETFIEDVGKEITLQAKPLDILIGVSEGMLA 419
 QY 421 ASLNAHKGVYCTTAHALEKTKYPSDYLKMKLEDKYHFSQCFADLFANMHTDFTTS 480
 DB 420 ASLNAHKGVYCTTAHALEKTKYPSDYLKMKLEDKYHFSQCFADLFANMHTDFTTS 479

QY 481 TFOEIASKDTVGQYESHRTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPPYTEKR 540
 DB 480 TFOEIASKDTVGQYESHRTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYPPYTEKR 539
 QY 541 RLKHFHEIEDLTYTKNEEHLCVLNDNRKPLIFMPRLDRVKNLGLVEMCKNKLK 600
 DB 540 RLTFHFEIEDLTFSDVNEEHLCVLNDNRKPLIFMPRLDRVKNLGLVEMCKNKLK 599
 QY 601 ELANLVVGGDRKRESKDLSEKAEKMFELIDKYNLNGOFWMISSQMNIRNVELRYI 660
 DB 600 ELVNLVVGGDRKRESKDLSEKAEKMFELIDKYNLNGOFWMISSQMNIRNVELRYI 659
 QY 661 CDTKGAFTVDPALYEAFTLVVEAMTCGLPTPATNGSPAEIIVHGKSGFNIDPYHQDA 720
 DB 660 ADTGAFVQPAFYEAFTLVVEAMSCGLPTPATNGSPAEIIVHGKSGFQIDPYHQDA 719
 QY 721 DILVDFEKKCKDPSHNDKISGGLKRIEKKYTKIYSERLLITGYGFPKHSNLER 780
 DB 720 DILADFEKKCKDPSHNEALSEGKRIOEKYTWQIYSDRLTLAAVYGFVKHVSKLDR 779
 QY 781 ESRRYLEMFYALKRYKLAESVPLAE 806
 DB 780 EIRRYLEMFYALKRYKLAOLVPLAE 805
 RESULT 10
 SUSL DAUCA STANDARD; PRT; 808 AA.
 ID SUSL DAUCA
 AC P49035;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sucrose synthase isoform I (EC 2.4.1.13) (Sucrose-UDP
 glucosyltransferase 1) (Susy-Dcl1).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulales; Apiales; Apiaceae; Apioidae; Scandiceae; Daucineae;
 OC Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nantaise;
 RX MEDLINE=95303982; PubMed=7784526;
 RA Sebkoova V., Unger C., Hardegger M., Sturm A.;
 RT "Biochemical, physiological, and molecular characterization of
 sucrose synthase from Daucus carota.";
 RL Plant Physiol. 108:75-83(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nantaise; TISSUE=leaf;
 RX MEDLINE=99178785; PubMed=10080700;
 RA Sturm A., Lienhard S., Schacht S., Hardegger M.;
 RT "Tissue-specific expression of two genes for sucrose synthase in
 carrot (Daucus carota L.).";
 RL Plant Mol. Biol. 33:349-360(1999).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- ENZYME REGULATION: Fructose acts as a noncompetitive inhibitor
 with an inhibition constant of 17.2 mM. In contrast, glucose
 inhibits uncompetitively with an inhibition constant of 4.3 mM.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- TISSUE SPECIFICITY: Expressed in stems, in roots at different
 developmental stages, and in flower buds, flowers and maturing
 seeds, with the highest levels in strong utilization sinks for
 sucrose such as growing stems and tap root tips.
 CC -1- SIMILARITY: Belongs to the glucosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X75332; CAA53081.1; -
 DR EMBL; Y16090; CAA53081.1; -
 DR PIR; S37560; S37560
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose_synth.
 DR Pfam; PF00534; Glyco_transf_1; 1.
 DR Pfam; PF00862; Sucrose_synth; 1.
 DR Transferrase; Glycosyltransferase; Multigene family.
 KM SEQUENCE 806 AA; 92474 MW; 1A40FCABAA4A5425 CRC64;

Query Match 80.1%; Score 3400.5; DB 1; Length 808;
 Best Local Similarity 79.1%; Pred. No. 2.1e-211;
 Matches 639; Conservative 76; Mismatches 90; Indels 3; Gaps 2;

QY 1 MAERALTIVHSLSRELDFTLHARNEILLALSRIEKGKGILOHIOIILEFAIPEENR 60
 DB 1 MEEPLTVHSLSRELDFTLHARNEILLALSRIEKGKGILOHIOIILEFAIPEENR 60
 QY 61 KL--ANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 118
 DB 61 KLDDHGNANAEVIKQEAIVSPWVALAIRPGVMEYIRVNVVALVEELTVAEYLHF 120
 QY 119 KEELVDGSSNGNVLDELDEPPNSSFPPRTLSKISNGVEFLNRHLSAKLPHDKESMPL 178
 DB 121 KEELVYSSDANFVELDEPAPFTASFPRTLSKISNGVEFLNRHLSAKLPHDKESMPL 180
 QY 119 LEFLVHCHKGKMMNDRIOMNLQHLARKEEVLGTLPEPTPAEENHFOELGLER 238
 DB 181 LEFLVHCHKGKMMNDRIOMNLQHLARKEEVLGTLPEPTPAEENHFOELGLER 240
 QY 239 GWDGTAERVLNEMIQLLDLLEATDPTCEKFLGRIPMVNVVILTPHGFADNVLYGYPD 298
 DB 241 GWDGTAERVLNEMIQLLDLLEATDPTCEKFLGRIPMVNVVILTPHGFADNVLYGYPD 300
 QY 299 TGGGVVYIIDQVRALENEMLRIKQGGNITRILITRLLPDAVGTTCGRLERVYGT 358
 DB 301 TGGGVVYIIDQVRALENEMLRIKQGGNITRILITRLLPDAVGTTCGRLERVYGT 360
 QY 359 HSDILRVPRTEKGIARVKSISFEKWPVLETYTEDVAHEISKEHLGTDLLIIGNXSDGN 418
 DB 361 HSDILRVPRTEKGIARVKSISFEKWPVLETYTEDVAHEISKEHLGTDLLIIGNXSDGN 419
 QY 419 IYASLLAHGLVTOCTIAHLEKTKYPSDSIYMKLEDEKHFSCOPTADLPAMNHTDFII 478
 DB 420 IYASLLAHGLVTOCTIAHLEKTKYPSDSIYMKLEDEKHFSCOPTADLPAMNHTDFII 479
 QY 479 TSTFOEINGSKDTVOYESHTAFTLGLRVVHGVIDVPDKENIVSPGADMEIYEPYTEE 538
 DB 480 TSTFOEINGSKDTVOYESHTAFTLGLRVVHGVIDVPDKENIVSPGADMEIYEPYTEE 539
 QY 539 KRRLGHFHEIEDLLYTKVNEEHLCVLNDKRNKPLFTMPRLDRYKNTGLVWKGCKRXP 598
 DB 540 KRRLGHFHEIEDLLYTKVNEEHLCVLNDKRNKPLFTMPRLDRYKNTGLVWKGCKRXP 599
 QY 599 LRELNLVVGDRRKESKDELEKAMKMFELIDRYNNGOFRWISSOMNIRVVELR 658
 DB 600 LRELNLVVGDRRKESKDELEKAMKMFELIDRYNNGOFRWISSOMNIRVVELR 659
 QY 659 YICDTKGAFOVALYFAGLVVEAMTCLPTFATCNGGPABIIVHSGSGFNIDPYHQD 718
 DB 660 YICDTKGAFOVALYFAGLVVEAMTCLPTFATCNGGPABIIVHSGSGFNIDPYHQD 719
 QY 719 AADILVDFEKKCKPSHWDKISOGGLRIEKKYTKWYISELLTLTGYYGKWKVSUNLE 778
 DB 720 AADILVDFEKKCKPSHWDKISOGGLRIEKKYTKWYISELLTLTGYYGKWKVSUNLE 779
 QY 779 RRESRYLEMFYALKYRKLAESVPLAE 806

DB 780 RLEIRRYLEMFYALKYRKLAESVPLAKD 807

RESULT 11

SUS2_TULGE STANDARD; PRT; 820 AA.
 ID SUS2_TULGE
 AC Q41607;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 OS Sucrose synthase 2 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 2).
 OS Tulipa gesneriana (Tulip).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Tulipa.
 NC NCBITaxID=13306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Apeldoorn;
 RA Balk P.A., de Boer A.D.;
 RL Submitted (Mar-1996) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC - CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC - SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X96939; CAA56440.1; -
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose_synth.
 DR Pfam; PF00534; Glyco_transf_1; 1.
 DR Pfam; PF00862; Sucrose_synth; 1.
 DR Transferrase; Glycosyltransferase; Multigene family.
 KM SEQUENCE 820 AA; 93015 MW; AA75AF8BD74A0EED CRC64;

Query Match 77.7%; Score 3301; DB 1; Length 820;
 Best Local Similarity 77.0%; Pred. No. 5.5e-205;
 Matches 620; Conservative 88; Mismatches 95; Indels 2; Gaps 2;

QY 1 MAERALTIVHSLSRELDFTLHARNEILLALSRIEKGKGILOHIOIILEFAIPEENR 59
 DB 1 MADRMATVHSLSRELDFTLHARNEILLALSRIEKGKGILOHIOIILEFAIPEENR 60
 QY 60 KKLANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 119
 DB 61 KKLANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 120
 QY 120 BELVUGSSNGNVLDELDEPPNSSFPPRTLSKISNGVEFLNRHLSAKLPHDKESMPL 179
 DB 121 BELVUGSSNGNVLDELDEPPNSSFPPRTLSKISNGVEFLNRHLSAKLPHDKESMPL 180
 QY 180 EFLRVHCHKGKMMNDRIOMNLQHLARKEEVLGTLPEPTPAEENHFOELGLER 239
 DB 181 EFLRVHCHKGKMMNDRIOMNLQHLARKEEVLGTLPEPTPAEENHFOELGLER 240
 QY 240 WGDGTAERVLNEMIQLLDLLEATDPTCEKFLGRIPMVNVVILTPHGFADNVLYGYPD 299
 DB 241 WGDGTAERVLNEMIQLLDLLEATDPTCEKFLGRIPMVNVVILTPHGFADNVLYGYPD 300
 QY 300 GGGVVYIIDQVRALENEMLRIKQGGNITRILITRLLPDAVGTTCGRLERVYGT 359
 DB 301 GGGVVYIIDQVRALENEMLRIKQGGNITRILITRLLPDAVGTTCGRLERVYGT 360
 QY 360 SDILRVPRTEKGIARVKSISFEKWPVLETYTEDVAHEISKEHLGTDLLIIGNXSDGN 419

Db 726 LLVDFDCKQAEPSWSKISGGGLRIEKKYTWKLYSRMLTMTGTVGFWKXVNSLBERE 785
 Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 Db 786 TRRYLEMYALKYRKMASTVPLAVE 810

RESULT 13
 SUS1_TULGE
 ID SUS1_TULGE STANDARD: PRT: 805 AA.

AC Q41608: 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sucrose synthase 1 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 1).
 OS Tulipa gesneriana (Tulip)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Tulipa.
 OC NCBI_TaxID=13306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Apeldoorn;
 RA Balk P.A., de Boer A.D.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. By non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X96938; CAA5639.1; -
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose_synth.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 DR Pfam; PF00862; Sucrose_synth; 1.
 KM Transferrase: Glycosyltransferase: Multigene family.
 SQ SEQUENCE 805 AA; 92254 MW; B749D8953C9338F0 CRC64;

Query Match 77.1%; Score 3273; DB 1; Length 805;
 Best Local Similarity 76.1%; Pred. No. 3.4e-203;
 Matches 613; Conservative 97; Mismatches 91; Indels 4; Gaps 3;

Qy 1 MAERALTIVHSGIRELDELTAHREIILALSRIGKGGKILQHIIIEPA-IPENR 59
 Db 1 MTERLTITIHISIKERIGDLSHHPNELALFSRFKQKMLERHQLTTEVESVPEADR 60
 Qy 60 KKLNGAFEEVLKASOEALIVLPWVALAVRPPGWETIRVNVHVLVVEELTVAEYLFK 119
 Db 61 EKLKQGVFEDTIRKSOEALIVLPWVALAIRPPGWETIRVNVHVLVVEE-CSIYLFK 118
 Qy 120 EELVDSGSGNFVLELDEFPNFSPPRTLSKISNGVEFLNRHLASAKLFDKESMPL 179
 Db 119 EDVDRSSQSNFVLEMFEPFANVPRLSLKISNGVQFLNRHLSSLFLDKESLYPL 178
 Qy 180 EFLRVHCKGKMMNDRIQNLALQHVLRKAEEYLGTLPTPECPAEFHRFOEIGLERG 239
 Db 179 NPLEHNHKGKTLMLNDRLQSLAQTLARKADRYLISIKDTPYSERHNSFOYGLRKG 238
 Qy 240 WGDPAERYLEMIQLDLLEATDPCTLEKFLGRIMVENVVILTPHGYEADNVLYGPD 299
 Db 239 WGDTRSRSENHLLDLLEAPDPTLEKFLGTMVENVVILSHGFAQNVLYGPD 298
 Qy 300 GGQVVYILDQVRALENEMLRKIQGLNITPRILITLDPVAGTTGQRLERKYGTEH 359

Db 299 GGQVVYILDQVRALETEMLKIKOQGLDITPRILITLDPVAGTTGQRLERKYGTEH 358
 Qy 360 SDLRVPRTEKGIYRKWISRFEEKWPLYETTEDVAHEISKEHLGTPDLLIGXSDGNI 419
 Db 359 THILRVPRIDKGLIRKWSRFE-WMPYLETYAEDVAHELGEMOATDILLIGNSDGNL 417
 Qy 420 VASLLAHRLGVTQCTIAHALEKTKYPSDSIYWKLEDKYHFSQPTADLFANHTDFTIT 479
 Db 418 VASLLAHRLGVTQCTIAHALEKTKYPSNSDIYAKFPDDQYHSCQPTADLLIMNGSDFIT 477
 Qy 480 STFOELASQDTVGQYESHTAFTLPGLRVVHGIDVPDPKKNISPGADMETIYPYTEK 539
 Db 478 STFOELASQDTVGQYESHTAFTLPGLRVVHGIDVPDPKKNISPGADMETIYPYTEK 537
 Qy 540 RLLKHFHEIDLLYTKYNEEHCVLNDKRPILFTMPRLDRVNLGLVEMGSKNKL 599
 Db 538 KRLTRLAHEIEELLYSVNEEHKFLVDKRNKPIIFSNARIDRVNMGVLELKGKDKL 597
 Qy 600 RELANLVVGGDRRKESKOLEEKAEMKMFELIDKYNLNGOFRWISSQMRIRNVELYRY 659
 Db 598 KELVNLVVADHGESKOLEEKAEMKMYKLIBEYKLGHIIRMSAQMNRVNGELYRY 657
 Qy 660 ICDTKGAFVQALYFAFGLTVVEAMTCGLPTPATCGGPAEIIYHKSNGFNIDPYHQA 719
 Db 658 IADTKGAFVQAFYEAFFGLTVVESWTCGLPTPATCGGPAEIIYHGVSGFHIDPYHQA 717
 Qy 720 ADIYVDFEKKKQPSHMDKISGGKRIEKKYTWKLYSRMLTMTGTVGFWKXVNSLBER 779
 Db 718 SEQLVSFEKKKEPDAHWEKISGGGLQRIYERKTYWKLYSRMLTLAGVGFWMKXVNSLDR 777

Qy 780 RESRRYLEMFYALKYRKLAESVPLA 804
 Db 778 TRRYLEMYALKYRKMASTVPLA 802

RESULT 14
 SUS2_ORYSA
 ID SUS2_ORYSA STANDARD: PRT: 816 AA.

AC P31924:
 DT 01-UTL-1993 (Rel. 26, Created)
 DT 01-UTL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase 2 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 2).
 GN SS2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eubatrachideae; Oryzae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Tainong 67;
 RX MEDLINE=92119223; PubMed=1531032;
 RA Yu W.P., Wang A.Y., Huang R.H., Sung H.Y., Su J.C.;
 RT "Isolation and sequences of rice sucrose synthase cDNA and genomic
 DNA.";
 RL Plant Mol. Biol. 18:139-142(1992).
 RN [2]
 RP SEQUENCE OF 605-816 FROM N.A.
 RX MEDLINE=92288314; PubMed=1534703;
 RA Wang A.Y., Yu W.P., Huang R.H., Huang J.W., Sung H.Y., Su J.C.;
 RT "Presence of three rice sucrose synthase genes as revealed by cloning
 and sequencing of cDNA.";
 RL Plant Mol. Biol. 18:1191-1194(1992).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

Db 243 RVLDTLHLLDLLEAPDPALEKFLGTIPMMFNVLSPHGYFAQSNVLGYPDITGGQVYV 302
QY 306 ILDOVRALLENMLARIKQOGLNTPRILITRLLPDAVGTTCGRLEKRYGTEHSDILRV 365
Db 303 ILDOVRALLENMLARIKQOGLDITPKILITRLLPDAAGTTCGRLEKRYGTEHSDILRV 362
QY 366 PPRTEKGIVRKWIISFEKVPWPLETYTEDVAHEISKEHAGTPDLLIGNXSDGNIVASLLA 425
Db 363 PPRNENGLRKWKISRFD-VMPYLETYTEDVSEIMKEMQAKPDLLIGNYSDGNIVATLLA 421
QY 426 HKLGVTCTIAHALEKTYPPDSDIYWKLEDEKYPSCQFTADLFAMNHTDFIITSTFOEI 485
Db 422 HKLGVTCTIAHALEKTYPPNSDIYLDKFDSDQYHFSQFTADLIAMNHTDFIITSTFOEI 481
QY 486 AGSKDTGQYESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEKRLKHF 545
Db 482 AGSKDTGQYESHAFITLPGLYRVVHGIDVDPKFNIVSPGADMSVYYPYTEDKRLTAF 541
QY 546 HPEIEDLLYTKVENEHLCVLNDNRKPIITWPRLDVKNLTGLVEMCGKNPKLRELANTL 605
Db 542 HPEIEELIYSDVENSEHKEFVLKDKKKPIIFSMARLDRYKMTGLVEMYGNARLRELANTL 601
QY 606 VVVGDRRKESKDLSEKAKEMKMFELIDKYNLNGQFRWISSQMRIRNVELYRYICDTKG 665
Db 602 VIVAGDHGKESKDRREQALEPKMYSLIDYKLGKHIRWISAQMNRRVNGELYRYICDTKG 661
QY 666 AFVQPALYEAFLTYVEAMTCGLPTFATCNGGPAEIIYHSGSGFNIDPYHGDQADILVD 725
Db 662 AFVQPAFYEAFLTYIESMTGCLPTIATCHGGPAEIIYDGVSGLHIDPYHSDKADILVN 721
QY 726 FFEKCKDPSHMDKISQGLKRIEKEYTKIYSERLTLTGYYGFWKHYNSLNERESRY 785
Db 722 FFDKCKADPSYWEISQGLQRIYKEYTKLYSERLMTLTGYYGFWKHYNSLNERETRY 781
QY 786 LEMFYALKYRKLAESVPLA 804
Db 782 IEMFYALKYRSLSAQVPLS 800

Search completed: June 2, 2004, 14:51:35
Job time : 21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 23:27:21 ; Search time 214 Seconds
(without alignments)
6807.225 Million cell updates/sec

Title: US-10-003-405-1

Sequence: 1 atgctcgaagcgtgctctcac.....tcaaaaaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB_seg.*
2: /cgn2_6/prodata/2/ina/5B_COMB_seg.*
3: /cgn2_6/prodata/2/ina/5A_COMB_seg.*
4: /cgn2_6/prodata/2/ina/5B_COMB_seg.*
5: /cgn2_6/prodata/2/ina/PTUTS_COMB_seg.*
6: /cgn2_6/prodata/2/ina/backfillseg_seg.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Query					ID	Description
No.	Score	Query	Length	DB		
1	147.8	57.1	3103	4	US-09-558-401C-57	Sequence 57, Appl
2	1053.4	40.1	2563	2	US-08-553-436A-7	Sequence 7, Appl
3	491.6	18.7	2700	4	US-08-684-005-1	Sequence 1, Appl
4	235.8	9.0	532	4	US-09-558-401C-56	Sequence 56, Appl
5	165.8	6.3	6386	2	US-08-483-376-1	Sequence 1, Appl
6	142	5.2	271	4	US-09-313-294A-1826	Sequence 1826, Appl
7	136.2	5.2	272	4	US-09-313-294A-1221	Sequence 1221, Appl
8	85.6	3.3	265	4	US-09-313-294A-1862	Sequence 1862, Appl
9	68	2.6	2236	4	US-09-697-367-1	Sequence 1, Appl
10	59.4	2.3	3625	1	US-08-356-354-3	Sequence 3, Appl
11	59.4	2.3	3625	2	US-08-778-656-3	Sequence 3, Appl
12	59.4	2.3	3740	1	US-08-356-354-1	Sequence 1, Appl
13	59.4	2.2	3740	2	US-08-778-656-1	Sequence 1, Appl
14	57.8	2.2	2930	1	US-08-356-354-5	Sequence 5, Appl
15	57.8	2.2	2930	2	US-08-778-656-5	Sequence 5, Appl
16	57.6	2.2	282	4	US-09-313-294A-427	Sequence 427, Appl
17	53.4	2.0	84	4	US-09-313-294A-7580	Sequence 7580, Appl
18	49.8	1.9	3509	1	US-08-175-471-6	Sequence 6, Appl
19	49.8	1.9	3509	3	US-09-078-862-2	Sequence 2, Appl
20	49.8	1.9	3509	4	US-09-866-153-12	Sequence 12, Appl
21	49.8	1.9	3509	4	US-09-693-467A-12	Sequence 12, Appl
22	48.2	1.8	3509	2	US-08-718-777-6	Sequence 6, Appl
23	48.2	1.8	3509	3	US-09-051-341-6	Sequence 6, Appl
24	45.4	1.7	1786	4	US-08-956-171E-618	Sequence 618, Appl
25	45.2	1.7	1445	4	US-09-697-367-19	Sequence 19, Appl
26	45	1.7	3539	3	US-08-853-948B-1	Sequence 1, Appl
27	44.8	1.7	3635	2	US-08-553-436A-5	Sequence 5, Appl

28	44.4	1.7	3509	2	US-08-429-054A-12	Sequence 12, Appl
29	44.2	1.7	210	3	US-09-125-984-1	Sequence 1, Appl
30	43.6	1.7	7218	1	US-08-332-463-14	Sequence 14, Appl
31	42	1.6	3534	4	US-09-697-367-15	Sequence 15, Appl
32	40.8	1.6	909	4	US-09-134-001C-334	Sequence 334, Appl
33	39.4	1.5	1664976	4	US-08-916-421B-1	Sequence 1, Appl
34	37.6	1.4	1158	4	US-09-943-681A-3266	Sequence 3266, Appl
35	37.4	1.4	1664576	4	US-08-916-421B-1	Sequence 1, Appl
36	37.2	1.4	339	4	US-09-621-976-8976	Sequence 8976, Appl
37	37	1.4	1336	4	US-09-328-352-4064	Sequence 4064, Appl
38	36.8	1.4	8961	4	US-10-304-708-80	Sequence 80, Appl
39	36.6	1.4	2156	3	US-08-965-762-16	Sequence 16, Appl
40	36.6	1.4	2156	4	US-09-911-927-16	Sequence 16, Appl
41	36.6	1.4	2156	4	US-09-911-927-16	Sequence 16, Appl
42	36.6	1.4	2156	4	US-09-911-882-18	Sequence 18, Appl
43	36.6	1.4	2156	4	US-09-911-882-18	Sequence 18, Appl
44	36.6	1.4	2156	4	US-09-911-888-16	Sequence 16, Appl
45	36.6	1.4	2156	4	US-09-911-888-18	Sequence 18, Appl

ALIGNMENTS

```

RESULT 1
US-09-598-401C-57
Sequence 57, Application US/09598401C
Patent No. 6596925
GENERAL INFORMATION:
APPLICANT: Perera, J. Ranjan
APPLICANT: Bagleton, Clare
APPLICANT: Rice, Stephen J.
TITLE OF INVENTION: Compositions and Methods for the
FILE REFERENCE: 11000.1036c2
CURRENT APPLICATION NUMBER: US/09/598,401C
CURRENT FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: PCT/NZ00/00018
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
PRIOR FILING DATE: 1999-03-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57
LENGTH: 3103
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-09-598-401C-57

Query Match      57.1%; Score 1497.8; DB 4; Length 3103;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1

QY      1  ATGGCTGACGCGTCTCTCACTCGCGTCCACAGTCTCCGTGAGCGCTTGGATGACCCCTT 60
DB      447  ATGGCTGATCGCATGTTGACTCGAAGCCACAGCCTTCGCGAGCGTTGGACGACCTTC 506

QY      61  CTTCGTCACAGGAAGAGATTTTGGCTTGTCTCAAGATCGAGGCGCAAGGAAAAGGA 120
DB      507  TCTGCTCACCGCAAGATATTGTGTGCTTCTTCCTTCAAGGTTGAAGCGCAAGGCGAAAGGC 566

QY      121  ATTTCGCAACACCATCAATATTCTAGAGTTGAAGCTATCCCTGAAGAGACAGAAAG 180
DB      567  ATCTTGCAACCGCACCAAGATTTTCTGATGTTGAGGCCATCTCTGAGAGACAGACGA 628

QY      181  AAGCTGCTAAATGATGCAATTTTGAATAATTAAGGTAAGTCAAGAGAGCATGTCTTTG 244
DB      627  AAGCTTCTTGAATGGGCGCTTTGGTGAATCTCTCAAAATCCATCAGGAAGCGATTGTCTG 686

QY      241  CCTGCATGGGTTGCACTTGCTGTTGTGTCGAAGGCGCGTGTTTGGAGATTCATTAGAGTG 300
DB      687  CCTGCATGGGTTGCTTGTGCTGTTGTGTCGAAGGCGCGGCGGTGTGGAGCACATCCGTGTG 746

```

QY	301	AAAGTTCACGGCCCTTGTGTTGAGGAACTCACTGTGTGTAGATATCTCATCTTCAAGGAA	360
Db	747	AAAGTTCACGGCCCTTGTGTTGAGGAACTCACTGTGTGTAGATATCTCATCTTCAAGGAA	806
QY	361	GAGCTTGTGTGAGAAAGTTCAAAATGGAACCTTGTGTTGGAATTTGATTTGAGCCCTTC	420
Db	807	GAGCTTGTGTGAGAAAGTTCAAAATGGAACCTTGTGTTGGAATTTGATTTGAGCCCTTC	866
QY	421	AACTCATATTCCTCCCGCCCAACTCTTCAAAATCCAAATTTGATATGTGTGAGAGTTCCTA	480
Db	867	AACTCATATTCCTCCCGCCCAACTCTTCAAAATCCAAATTTGATATGTGTGAGAGTTCCTA	926
QY	481	AATGTCACCTTTCGGCAAAATTTGTTCCATGACAGAAGACATGACCCCTTGTGCTGAA	540
Db	927	AATGTCACCTTTCGGCAAAATTTGTTCCATGACAGAAGACATGACCCCTTGTGCTGAA	986
QY	541	TTCTCTCAGAGTTCATTGTCAACAAGGGCAAAACATGATGTTGTAATGACAGAAATTCAGAC	600
Db	987	TTCTCTCAGAGTTCATTGTCTACACAAGGGGAAAGAACATGATGTGTAATGCCAGATCCAGAA	1046
QY	601	TTGAATGCTCTTCAACATGTTTGTGAGAAAGAGAGAGATATCTGTGAATCCCTACCTCTCT	660
Db	1047	GTGTTCTCTCTCAACATGTTCTGTGAGAAAGGGGAGAGATATCTGACTCTGCTCAAAACC	1106
QY	661	GAGACACCAGTATGCCGAAATTCGAAACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG	720
Db	1107	GAGACACCAGTATGCCGAAATTCGAAACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG	1166
QY	721	GGTGCACCCGAGAAACGGGTGTCTGAGATATGATTCCACTCTTTTGGATCTTTGAGGCA	780
Db	1167	GGTGCACCGGTGAGCCGGCTCTGAGATATGATTCCACTCTGTGAGATCTCTTGAAGCT	1228
QY	781	ACTGATCTCTTACACCTCTTGAGAAAGTCTCTTGGGAGAAATCCCATGATGTTCAATGTTGTG	840
Db	1227	CCGACCCGCTGACCTCTGAGAAAGTCTTGGATGAGGTTCCCATGATGTTCAATGTTGTG	1286
QY	841	AATTCACCTCCCGACGGATCTTGCTCAAGACAAATGTTTGGGATATCCGACACCGGT	900
Db	1287	AATCATGCTCCCGACGGATCTTGCTCAAGACAAATGTTTGGGATATCCGATACCGGT	1346
QY	901	GGGACGGTGTGTTACATTTGGATCAAGTCGAGCTTGGAGAAATGAGATGCTCTCCGT	960
Db	1347	GGCAGGTTGTTTACATCTTGATCAAGTTGCTGAGACGAGCTCTGTGGTATCCGATACCGGT	1406
QY	961	ATTAAGCAACAAGGACTGACATCAACATCCCTCGAATCTCATTTATCTAGACTTCTTCT	1020
Db	1407	ATTAAGCAACAAGGACTGATATTAATCTCTCGAATTTCTCATTTATCACTCGGCTTCTTCCA	1466
QY	1021	GATGCTGTCCGAAACAATGCGGTCAAGAATCTTGGAAAGATATACGGAACAGAGACTCG	1080
Db	1467	GACGCGGTTTGAAACCACTGTGTGCGAGGCTTGGAAAGTTTTGGGACGAGATCTCC	1528
QY	1081	GATATTTCTTCAGTACCTCTTCGAAACAGAAAGGGAATTTGTTCGAAATGATCTCAAGA	1140
Db	1527	CACATTTCTTGGGTCTCCCTTCGAAATAGAGAGGAGTGTCTCCGAAATGTATTTCCGG	1586
QY	1141	TTTGGAAAAAGTGTGGCCATATCTTGGAAACCTACACAGAGATGTTGCTCATGAATCTTCC	1200
Db	1587	TTTGG---AGGTGTGGCCCTTATTTGGAAAGATACATGTGAGATGTGTCCAGAGGAATCTGCT	1643
QY	1201	AAAGGTTGACCGGACGCGACAGATGTGATCATCGGAAACAGGCAAGGCAATATCGTCT	1260
Db	1644	GAGAGTTTGCAGGGGCAAGCTGATTTGATCATCGGAAACCTACAGTGAATGGAACATTTGTT	1703
QY	1261	GCCTCTTGTCTCGACATPAATTTAGTGTCAACACAGTGCACCATGCGCCATGTCTTGGAG	1320
Db	1704	GCTTCTTGTTAGCAATPAATTTAGTGTACACAGTGAACAAATGTGCCATGTCCCTCGAG	1763
QY	1321	AAAGCAAAATATCCGATTTCCAGATATCTATTTGGAAGAGCTTGAAGACAAATACATTTCC	1380
Db	1764	AAAGCAAAATATCCGAGTCAAGACATATATCTGGAAGAAATTTGAGAAAGATACACATTC	1823

OY	1381	TCCTGCCAATTTAGAGCGATACCTTTTGGCATAGAACCATACAGATTTTCATATCACCAGT	1440
Db	1824	TCCTGCCAATTTCACTGCTGATCTCATGCGCATGAAACACACCGACTTCTATATCACAGC	1883
OY	1441	ACTTTCAGGAAATTTGACGAAAGCAAGACACATGTTGTCATATACAGAGCCACACTGCT	1500
Db	1884	ACCTTCCAGAAATTTGCTGGAAAGCAAGATACAGTGGGGCAGTATGAGAGTCATGAC	1943
OY	1501	TTTCACTCTCTCGGTCTCTACCGGTGTGTACATGATATGCAATGTGTGATATCCCAATTC	1560
Db	1944	TTCACTCTCTCTGAGACTCTACCGAGTGTGTCCAGGGATCCAGAGCTTTCGACCCGAAAGTTC	2003
OY	1561	AACATTTGTTCCCTGAGTGTGATATGAGATATATCTCTCTTACACCGAAGACGCG	1620
Db	2004	AAATTTGTTCCACGAGTGTGTGATATGAGATATATCTTGTGTTACCCGAACAGAGCGG	2063
OY	1621	AGTTTGAAGCATTTCCATCCTGAGATGGAAGACCTTCTTATACACCAAGTTGAGATGAA	1680
Db	2064	CGTTTGAATTCCTTCCACCTCGATATGAGGAACTCTCTTCAGCGATGTTTGAACAAAG	2123
OY	1681	GAACACTATATGTGTCTCAATGACCGGACCAAGCCCAATTCGTTCACATGCGAAGCTT	1740
Db	2124	GAACTCTTGTGTGTGTGAAAGATTAAGAAAGAGCTTATTTTTCACCATGCGAAGCTG	2183
OY	1741	GATCGTGTCAAGAACTTAAACCGAGCTCGTGAAGTGTGCGGACAGAAACCCAAAGTTCGT	1800
Db	2184	GACCGTGTCAAGAACTTGAACAGGCTGTGTGAGGTATGCGCAAGAACTCCAAAGTTGAG	2243
OY	1801	GAGTTTGCCTAACCTCGATATGTTATGATGTGTGATAGGCGAAAGAAATCTAAAGTTTGGAA	1860
Db	2244	GAACTCGCAACTTGTGTGTGTGTGAGGTGAACGAGAAAGAAAGTTGAGAGAACTTGGAA	2303
OY	1861	GAGAAAGCTGAAATGAGAAAGAAATGTTTGAAGCTGATCGACAAATGACCAACTTGAACGCGCAA	1920
Db	2304	GAGCAAGCTGATGATGAAAGAAATGTATAGCACTTACCGAAAGATCAAGCTGAATATGCGCAG	2363
OY	1921	TTTCAGATGATATCATCTCAATGAAACAGATCCGAAATGTTGAATTTTACCGATACATT	1980
Db	2364	TTTCAGGTGATATTCCTCCAGATTAACCGGATGAGAAATGAGAGCTCTACCGGTACATC	2423
OY	1981	TGCAACACGAAAGGTGCTTGTATGACAGCTGCACTGTATGAAGGCTTGTGATGACAGTT	2040
Db	2424	TGTATCAAGAAAGGAGCTTCGTTCACACCGCTATCTATGAAGCTTTCGGTTTACCGTG	2483
OY	2041	GTGAGAGCAATGACTTGTGCGGTTTGGCCAACTTGGCAACTGTAAACGGTGAACACGCCAG	2100
Db	2484	GTTGAGGCGCAATGACTTGTGATGTCCCAACTTGGCACTTGGCAATGGTGAACACAGCTGAG	2543
OY	2101	ATTATTTTGTCCATGAGGAATCTGTTTCAACATGATCCTTACACATGGATCAAGCTGT	2160
Db	2544	ATCATTTGTGACATGCAATCGGGCTACACATATATCTTACATGGTGAACAGCGGCGC	2603
OY	2161	GACATTACTCGTCGATTTCTTTGAAAGGTATGAAGAAAGTCCATCTCACTGCGATTAAGATC	2220
Db	2604	GAGCTTCTTGTAGACTTCTTTCACAAAGTGCAAGATTTGACAGTCCACTGCGAGCAAGATC	2663
OY	2221	TCCCAAGAGGCTTGAACAACGAATAGAGAGAAAGTATACATGGAAGATTTTACTCGGAGGA	2280
Db	2664	TCAAAAGGTGTCCATGCAAGAAATTTGAAGAAAGATACATGGAATATATTTCTGAGAGG	2723
OY	2281	CTATTTGACCCGTGACAGGAGTGTATGGAATTCGGAAGCATGTTTTCACACTTGAACGCGGT	2340
Db	2724	CTGTTTGAACCTGACTGCGGTGTATAGGCTTCTGGAAGCATGTGACTTAACCTTGAATCGGCGC	2783
OY	2341	GAGAGTCGTGCTTACCTTGAAGATGTTTATGTCTTAAATACCGTAAGCTGGCTGAATCA	2400
Db	2784	GAGAGTCGCGCGTACTTGAATATTTCTATGCTCCCTCAAGTATCGCCCACTGGCAAGCTCT	2843
OY	2401	GTTTCATTTGCAAGAGACTAA 2422	
Db	2844	GTTCTCCGCGTGTGAGATAA 2865	


```

RESULT 2
US-08-553-436A-7
Sequence 7, Application US/08553436A
Patent No. 5866790
GENERAL INFORMATION:
APPLICANT: HESSE, Holger
APPLICANT: KULLER-ROBER, Bernd
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostroienk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,436A
FILING DATE: 17-NOV-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/01671
FILING DATE: 20-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 4317596.1
FILING DATE: 24-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mellman, Edward
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: F/951-117
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2563 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Beta vulgaris
INDIVIDUAL ISOLATE: Saccharoseynthase
IMMEDIATE SOURCE:
LIBRARY: phage lamda zap
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2303
US-08-553-436A-7

Query Match 40.1%; Score 1053.4; DB 2; Length 2563;
Beet Local Similarity 67.8%; Pred. No. 0;
Matches 1535; Conservative 0; Mismatches 717; Indels 12; Gaps

Dy 164 CTGAAGAAACAGAAAGAAAGCTCGCTAATGTCGATTTTGTGAATATGAAGCTAGTC 223
|||
Db 1 CTGCAGAGAGGAAACAAATCTTACGATGCGCCCTTTAGCAATCTTAGTGTGCTG 60

Dy 224 AGAAGCATGCTGTTCCTTCAGTGGATTGCACTTGCTGTCGAAGCCCTGATGTTT 283
|||
Db 61 AGAAGCAATAGTTTCTCCCTTGTGTTGTAATAGCAAGTCGTCGAAGACTGAGATT 120

Dy 284 GGGAGTACTTATGATGAAATGTTACAGCCCTTGTGTTGAGAAATCACTGTTGCTGAGT 343
|||
Db 121 GGGATATATGTTGCTGTTTATATGCTCTGAACTGAATGTGAGAGAGCTAATCTGTCTGAGT 180

```

Oy	34	ATCTCCACTTCAAGGAGAGCTGTGATGAGAACTTGAATGAAACTTTGTTGGAT	403
Db	181	ATCTCCATTTCAAGGAAAGACTGTGATGAGAAAGCTGATGACCACTATGTCTTGAGC	240
Oy	404	TGATATTTGAGCCCTTCAACTCATCATATCCCGCCCAACTCTTCAAAATCCATTTGTA	463
Db	241	TTGATTTTGAGCCCTTTAAATGAAATCAAGTCCACGCTCCAACTCGCTTCAATCAATTTGTA	300
Oy	464	ATGATGTGAGATTCTTAAATGTCACTTTCCGAAATTTGTTCAATGACAGAGAGCA	523
Db	301	ATGATGTTCAGTTCCTCAATCGGACCTGTATCAACAGATTTTCGACAACTATGCT	360
Oy	524	TGCACCCCTTGTCTGGAATTCCTGAGAGCCATTGTCAAGGGCGAAGAACATGATTTGA	583
Db	361	TGAGGCCGTTACTGATTTTCTTGATGTGCAAAACATAAGAGTTGTCAATATTTGA	420
Oy	584	ATGACAGAAATTCAGAACTTGATATGCTCTTCAACATGTTTGAAGAAACAGAGAGTATC	643
Db	421	ATGATCGGATTCAGACTATTCAGAGGTCTTCAGTCTGCAATGTCTTAAAGCTGAGATATTC	480
Oy	644	TTTGTATACCTTACCTCTCGAGACACCAATGTCCGAAATTCGAACACCGGTTCCAGAAATCG	703
Db	481	TTATCAAACTTCCAGCAGATACACTTACTGTGAGTGGAAATTTGTAATCCAAAGTATGG	540
Oy	704	GTTTGAAAGAGTTTGGGGTGAACCCGACAGAACCGTCTCGAGATATCCAACTCTTTT	763
Db	541	GTTTGAAGAGGGCTGGGGTGAATCTGTGAAGAGGTTCTAGAAATATGATCTTACTAC	600
Oy	764	TGATCTTTCTTGAGGCAACTGATCTTTGCACTCTTGAGAACTCTTGGGAAATCCCA	823
Db	601	TGATATCTCTTCAGGCTCCGATCCGTCATATTAGAGACATTTCTGGGAAGACTTCCCA	660
Oy	824	TGAGTTCAATGTTGTGATCTCACTCCCAACGGATCTTGCTCAAGACATGTGTTTGG	883
Db	661	TGATGTTTAAATGTGTCAATTTTGTCTGACATGATATTTTGGACAGGCAATGTCTCG	720
Oy	884	GGTATCCCGAACCCGGTGGGCAAGTTGTTTACATCTTGGATCAAGTCGAGCTTGGAGA	943
Db	721	GCTTGCCGACACTGGTGGGCAAGATATTATATTACTTGACCAAGTCCGGCTCTGGAAC	780
Oy	944	ATGAGATGTCCTCCGATTAAGAACAAAGACCTCAACATCAACCCCTCGAATCTCTCATTA	1003
Db	781	ATGAAATGCTTCCAAAGATTAAGAAAGACATGATGTGATCTCTAGAAATTTCTTATCG	840
Oy	1004	TTACTAGACTTCTTCCATGATGCTGTCCGAAACAATGCGGTCAACGACTTGAGAAAGTAT	1063
Db	841	TGATTCGGTTGATTCCTGACCGCTTAAAGGACACGTGCATCAACGATATGAGAAAGTCA	900
Oy	1064	ACGGAACAAGCACTCGGATATTTCTTGAGTACCTTTCAGAACGAAAGGAAATTTGTT	1123
Db	901	GTGGAACAAGCACTGATGATCTCTGAGAGTTCTTTTCCGATCAGAGAAAGAAATCCCTCC	960
Oy	1124	GAAATGGAATCTCAAGATTTGAAAAGATCGGACATATCTTGGAAACCTACACAGAGATG	1183
Db	961	GCAATGGAATCTGATGATTTG--ATGATAGCTTTATTTAGACCTTCACTGAGATG	1017
Oy	1184	TTGCTCATGAATCTCCAAAGAGTTGACAGCGCACGCCAGATCTGATCATCGGAAACACAA	1243
Db	1018	CAGCTGTGTAATATTGCGAGATTTGACAGGATCGTCCAGATCTGATTAATTTGGCAACTTACA	1077
Oy	1244	GCGACGGCAATATGCTGTGCTCTTGTCTCGACATTAATTAAGGTGTACACAGTGACCA	1303
Db	1078	GCGATGGGAATATGTTGTCTTTATTTATGTCACAAAATGGGGTCAACCCAGTGACATA	1137
Oy	1304	TCGCCCATGCTTTGAGAGAGCAAAATATCCAGATTCAGATATCTATTTGGAAGAACCTTG	1363
Db	1138	TAGCCCATGATTTGAGAAACCAAGTATCCAGATTTCTGATATTTACTGGAAGAAATTTG	1197
Oy	1364	AAGCAAAATACATTTTCTTTGCAATTTTACAGTGAATCTTTTGGCAATGAAGCATTCAG	1423
Db	1198	AGGCAAAATACATCTTCTGTGTAAATTTTACGTGACTGTGATGGCAATGATATGTGTG	1257

QY 1424 ATTTCATCATCAACGATACCTTCCAGAAATTCGAGAGACAGACCTGTTGCTCAT 1483
 DB 1258 ATTTTCATCATTAACGATCTTACCAAGAGATAGCTGAGAACAGAAATCTGTTGCTCAT 1317
 QY 1484 ACGAGGACACACGCTTCTCTCTGCTCTACCGGTTGTCATGATGATGATG 1543
 DB 1318 ATGGAAGCATTAAGCTTCTCTCTGCTCTACCGGTTGTCATGATGATGATG 1377
 QY 1544 TGTTCATCCCAATTCACATTTGTTCCCTGCTGATGATGATGATGATGATG 1603
 DB 1378 TCTTTGATCCCAAGTTTAATTTGCTGCGAGGGGAGACATGCGCATCTACTCCAT 1437
 QY 1604 ACACCGAAGAGAGCGGA---GTTGAAGCATTTCCATCTCGAGATGAGACCTTCTT 1650
 DB 1438 TTTCAAGAGAGAGATGCTACTGCTCACTTCACTTCACTTATAGAGACGCTCCAT 1497
 QY 1661 ACACCAAGTTGAAGATGAAGAACACTATGTCGCTGAAGACCGCAAGCAACATTC 1720
 DB 1498 TCAAACTGACGACGAGAACAGAAACATTTGTTGATGATGATGATGATGATG 1557
 QY 1721 TGTTCACATGCAAGGCTGATGCTGCTCAAGAACTTAACCGGACTCGTCAAGTGTGCG 1780
 DB 1558 TATTTTCATGCGCAGGCTGACGCTGAGAGAAATTAACAGGGCTGATGATGATG 1617
 QY 1781 GCAAGAACCAAGTTGCTGATGCTGCTCAACCTGCTGATGATGATGATGATG 1837
 DB 1618 GCAAGAACCAAGTTGCTGATGCTGCTCAACCTGCTGATGATGATGATGATG 1677
 QY 1838 GAAAGGATCTTAAGATTTGGAAGAGAGGCTGAATGAAGAAATGTTGATGATG 1897
 DB 1678 TAAAGGATCTTAAGATTTGGAAGAGAGGCTGAATGAAGAAATGTTGATGATG 1737
 QY 1898 ACAAGTACATGCAAGGCTGATGCTGCTCAAGATGATGATGATGATGATGATG 1957
 DB 1738 AGAGATTAATTTAAGAGACATTTGCTGATGATGATGATGATGATGATGATG 1797
 QY 1958 ATGTTGAACCTTACCGATACATTTGCGACAGAAAGGCTGCTGATGATGATG 2017
 DB 1798 ATGTTGAACCTTACCGATACATTTGCGACAGAAAGGCTGATGATGATGATG 1857
 QY 2018 ATGAAGCTTGAAGTGAAGTGTGAGGAGCAATGATGATGATGATGATGATG 2077
 DB 1858 ATGAAGCTTGAAGTGAAGTGTGAGGAGCAATGATGATGATGATGATGATG 1917
 QY 2078 CCTGTAACGCTGACCAAGCAGATTAATGCTCAATGAGAAATCTGTTCAACATGATC 2137
 DB 1918 CCTGCAACGCTGACCAAGCAGATTAATGAGAAAGGCTTCAAGATTTCAATGATC 1977
 QY 2138 CTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2197
 DB 1978 CATATCATGCTGATCA---GGCAGAAAAAATGATGATGATGATGATGATG 2034
 QY 2198 ATCATCATGCTGAGATGAAGTCTCCCAAGAGGCTGAAACGAATGAGAGAGATTA 2257
 DB 2035 ATCCAACTACTGAGCTTAATTTCTGCAAGAGGCTTCTTAAGATTAAGAAAGATTA 2094
 QY 2258 CATGAAGATTTACTCGAGAGACTATGACCTGACAGAGATGATGATGATGATG 2317
 DB 2095 CTTGCAAAATTAATTTGAAAGTTAATGATGATGATGATGATGATGATGATG 2154
 QY 2318 ATGTTTCAACCTTGAAGCGCGTGAAGTGTCTGTTACCTTGAAGATGTTTATGCTT 2377
 DB 2155 ATGTTTCAACCTTGAAGCGCGTGAAGTGTCTGTTACCTTGAAGATGTTTATGCTT 2214
 QY 2378 AGTACCGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2421
 DB 2215 AGTTCGATGATGCTGCAACTCTGTTCCGCTGCAACAGATGA 2258

RESULT 3
 US-08-684-005-1
 ; Sequence 1, Application US/08684005
 ; Patent No. 6682918

; GENERAL INFORMATION:
 ; APPLICANT: Haselkorn, Robert
 ; APPLICANT: Buelkema, William J.
 ; APPLICANT: Bauer, Christopher C.
 ; TITLE OF INVENTION: BACTERIAL SUCROSE SYNTHASE
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: United States of America
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/684,005
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: ARSB:508
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (512) 474-7577
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2700 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-684-005-1
 ; Query Match 18.7%; Score 491.6; DB 4; Length 2700;
 ; Best Local Similarity 54.0%; Pred. No. 3.6e-146;
 ; Matches 1079; Conservative 0; Mismatches 905; Indels 15; Gaps 3;
 QY 378 TTCAATGAGAACTTTGTTGGAATGATTTGAGCCCTTCAACATCATTTCCCGC 437
 DB 607 TCCCTATGAAGGGGATATATAGAACTGATTTCCGCCCTTCAACATTAACCCCAAC 666
 QY 438 CCACACTTTTCAAAATCCATGCTGATGATGATGATGATGATGATGATGATG 497
 DB 667 CATCCGCGATCCAGAAATATGCGAAGGCTGATGATGATGATGATGATGATG 726
 QY 498 AAAATGCTTCATGACAAAGAGA---GCATGCAACCTTTGCTGATTTCTCAAGTCCA 554
 DB 727 TAAACTTTTCAAGACTCGCAACATGCTGGAAGCTGTTTAATTTCTTGGGCTTACA 786
 QY 555 TTGTCAACAGGCGCAAGAACTGATGTTGAATGACAGAAATTCAGAACTTGAATGCTCTTCA 614
 DB 787 TAATTACAAGGTAATCACTTAATTAACATCAATTAATCAACAGCAACAATATATC 846
 QY 615 ACAATGTTTGAAGAAACAGAGAGATCTTGTGATGATGATGATGATGATGATG 674
 DB 847 ACAGCAAGTTAABAAACCGCTTAATCTTTGAGAGATGATGATGATGATGATG 906
 QY 675 GCAATTCGAACACCGGTTCCAGAAATCGTTTGAAGAGATGATGATGATGATGATG 734
 DB 907 ACAATTCGCGCTGAATTAABAACTAATGATGATGATGATGATGATGATGATG 966
 QY 735 ACGGCTGCTGAAGATGATCAACTCTTTTGAATCTTTGAGGCAACTGATCTTGGAC 794
 DB 967 TCGTGTGCGGGAATCACTTAACATTTTGAATGATGATGATGATGATGATGATG 1026
 QY 795 CTTGAGAAATCTCTTGGAGAAATCCCATGATGATGATGATGATGATGATGATG 854
 DB 1027 CTTGAGAACTTTTATCTCTGCAATCCGATGATGATGATGATGATGATGATG 1086

QY 855 CGGATACCTGCTGAGCAATGTTTGGGGTATCCGACACCGGTGGGAGTTGTTA 914
 DB 1087 CGGTTGGTTGAGCAAGAGGGGGTTTAAAGTCTCCAGATCTAGTGTGTAAGTGTGTA 1146
 QY 915 CATCTGGATCAAGTCCGAGCTTTGAGAAATGAGATGC-----TCTCCGTATATA 965
 DB 1147 GGTCTTGAACCAAGCTAAGAAATTTAAGAAAGCACTGCAAGAAAGATGCATATCTGACG 1206
 QY 966 GCAACAAAGACTCAACATCACCCCTCCGATCTCTATTTATCTAGACTTTCTCTGATGC 1025
 DB 1207 TTTAGAGTATGAAACGTCGACCCAGAGTATTAATCTCAACCCGTCTGATTTCTTAATG 1266
 QY 1026 TGTGGAACAACATGCGGTGCAAGACTTGAAGAAATGAGAAACAGAGCACTGCAAT 1085
 DB 1267 TGACGAAACGTTTGTATCAAAAGGTTAAGAAAGTCTACGTAACAGAAACGCTGGAT 1326
 QY 1086 TCTTCGAGTACCTTTCAGAACAGAAAGGAATTTGCGAAATGATCTCAAGATTTGA 1145
 DB 1327 TTTGCGGTACCTCTGCGGGAGTTTAAACCAAGATGACGAAATTTGCTGATTT 1386
 QY 1146 AAAAGTCTGGCATATCTTGAACCTTACACAGAGATTTGCTCATTAATCTCCAAAGA 1205
 DB 1387 CGAGTTTGGCTTATCTAGAAACCTTTCATGACTCAGAAAGAAATTTGGCAGA 1446
 QY 1206 GTTGCAGCGGACGCGAGTCTGATCATGGAACACAGCGGCAATATGTCGCTC 1265
 DB 1447 ATTCAGAGTACAGCACTTAATCTGTGTAATTAATCTACGGAACCTTAGTGTCTT 1506
 QY 1266 CTTCCTGCAATTAATAGGTGTCAACAGTGCACATCCGCTTCTTGGAGAAAC 1325
 DB 1507 TCTGTGACGAGAGGAAAGTAACTTACCAATGAAACATGCTGATTTAGAAATC 1566
 QY 1326 AAAATATTCAGATTCATATCTATGGAAGAGTGAAGCAATTCATTTCTCTTG 1385
 DB 1567 CAATATCTTGTATGTAACCTTCTAGCAAGATTTGGAAGAAATATCTCTCTT 1626
 QY 1386 CCAATTTACAGCTATCTTTTGAAGTAAACATPACATTTCAATCAACAGTACTT 1445
 DB 1627 ACAATATCCGCTGATTAATAGCTATTAATGCTGTAATCTGCTACAGAGCACTTA 1686
 QY 1446 CCAGAAATTCAGAGAACAGACACTGTTGTCAATACAGAGCACTGCTTTTAC 1505
 DB 1687 TCAAGAAATTTTGGCACACAGACAGATAGGGGAGATAGTCTTAACAATGCTTAC 1746
 QY 1506 TCTTCTGTGCTTACCGGTGTGATCATGTATGATGTGTTGATCCCAATTCACAT 1565
 DB 1747 CATGCCGGAATCTATCATGTGTCAACGGCAATTAATTTAGCCCAATTTAAGT 1806
 QY 1566 TGTTCCTGCTGTGATGATGAGATATCTCCCTTACACCAAGAGAGCGAGCTT 1625
 DB 1807 TGTACCGCTGTGTGAATGAATTTCTACTTTCTTACACCAATCTCAAAACAGAT 1866
 QY 1626 GAAGCATTTCCATCTGATGAGAGACCTTCTTACACCAAGTGAAGTGAAGACA 1685
 DB 1867 AGAAAGCATGCGATGCGCTTAAGAGAAATGCTGTTTACCTTAAGAAATTTAGCAAT 1926
 QY 1686 CTATATGTGTCAATGACCGCAACAGCAATTTCTGTTCAATGCTCAAGGCTTATGCT 1745
 DB 1927 CTTGCGGCAATCTGACGACCAAAATPACGCTTATTTCTCAATGCGGCACTTGACCG 1986
 QY 1746 TGTCAAGAACTTAAACCGGACTGTGATGATGTGCGGCAAGAACCAAGTTCGGAATT 1805
 DB 1987 AATTAATAAACTCAAGGTTTGGCAGAAATGCTTGTGTAAGTCAAGAAATTTGCAAGCG 2046
 QY 1806 GGTAACTCTGATGTTAGGTGTGATAGCGAA---AGAAATCTAAAGATTTGGAAGA 1862
 DB 2047 TTGCAACTTAATTTTAAAGTGAAGTAACTGTGTAAGAAATGAGAAATTAAGAAAGA 2106
 QY 1863 GAAGGCTAAATGAAGAAATGTTGAGCTGATGACAAATTAATTTGAAACGCGCAATT 1922
 DB 2107 AAAAGACGAATTCGCAAACTTTACCGGATTAATGACGAATTAACCTGACATGCGCAAAAT 2166

QY 1923 CAGATGATATCATCTCAATGAACAGATCCGAATGTTGAACCTTACCGATACATTG 1982
 DB 2167 TCGCTGTAGGTGTGGCTTATCCAAATAATGATCCGGTGAATTTATGCGTCAATTTG 2226
 QY 1983 CGACAGAAAGGTGCTTTTGTACAGCTCGATTTGTATGAAGCTTTGATTAAGTTGT 2042
 DB 2227 CGATGCGCAAGGCAATTTTGTACAGCCAGCATTAATTTGAAGCCCTTTGGTGAACATCT 2286
 QY 2043 GGAGGCAATGACTGGGTTTGGCCACATTCGCAACCTGTAAAGGTGAGACCAAGAGAT 2102
 DB 2287 GAGTCAATGATTTCCGATTTGCCAATTTGCTTACCAATTTGGGGGCCATTTGGAGAT 2346
 QY 2103 TATTTCCATGGAATCTGTTCACATTTGATCTTACCATGATGATCAAGTGTGTA 2162
 DB 2347 TATTCAGATTAAGTTAATGCTTCTTACATTTAACCTTACATCTAGAAAGAACGCAAC 2406
 QY 2163 CATCTGCTGATTTCTTTGAAAGTGAAGAAAGATTCATCTGACGGATTAAGATCTC 2222
 DB 2407 AAAATTTCTGATTTCTGACCAATGCGAAACAAATCTTAATTTGGAACATTAATTTT 2466
 QY 2223 CCAAGAGGCTTGAAGAGATGAGAGAGTATACATGAAATTTTACTCGAGAGACT 2282
 DB 2467 CGAAGAAAGCATGACAGAGTATATGATACATACCTGGAATTAACACACATTAAGCT 2526
 QY 2283 ATTGACCTGACAGAGTATGATTTCTGGAAGCATGTTTCAACCTTGAACCGCGTGA 2342
 DB 2527 GTTAACCTTACCTGATTTACGGCTTCTGGAATTTTAACTCGAAGAAAGCGAAGA 2586
 QY 2343 GAGTGTGTTACTTTGAG 2361
 DB 2587 TTTATTAACGCTACCTTGA 2605

RESULT 4

US-09-598-401C-56
 ; Sequence 56, Application US/09598401C
 ; Patent No. 6596925
 ; GENERAL INFORMATION:
 ; APPLICANT: Perera, J. Ranjan
 ; APPLICANT: Eagleston, Claire
 ; TITLE OF INVENTION: Compositions and Methods for the
 ; FILE OF INVENTION: Modification of Gene Expression
 ; FILE REFERENCE: 11000.1036c2
 ; CURRENT APPLICATION NUMBER: US/09/598,401C
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: PCT/N200/00018
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
 ; PRIOR FILING DATE: 1999-03-25
 ; NUMBER OF SEQ ID NOS: 120
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 56
 ; LENGTH: 532
 ; TYPE: DNA
 ; ORGANISM: Pinus radiata
 ; US-09-598-401C-56

Query Match

9.0%; Score 235.8; DB 4; Length 532;
 Best Local Similarity 76.8%; Pred. No. 7,1e-65;
 Matches 288; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGCTAGAGCTGTCTCATCTGCGCTCACAGTCTCCGTGAGCGTTTGATGAGACCTT 60
 DB 158 ATGCTAGATCCAGATTTGATCGAAGCCACAGCTTGGCGGAGCTTTGAGACGACCTT 217
 QY 61 CTGCTACAGGAAGAGATTTTGGCTGTCTCAAGATCGAGGCGAAAGAAAGGA 120
 DB 218 TCTGCTACCGCAACGATTTGTGGCTCTCTTCAAGGTTGAAGCGCAAGGCGC 277
 QY 121 ATTCGACACACATCAATTAATTTAGAGTTGAAGTATCTCTGAAGAAAGAAAG 180

Db 278 ATCTTGACGCCCCACAGATTTTCTGAGTTTGAGGCCATCTGAGAGAGAGAGCA 337
Oy 181 AAGCTGCTAATGATGATTTTGAATATGAAGGCTAGTACGAGAGCATGCTGTG 240
Db 338 AAGCTTCTGATGAGGCTTTGCTGATGAGTCTCAATCATCTCAGAGAGCATTTGTG 397
Oy 241 CCTCATGAGGTTGACATTTCTGTTCTGTCAGAGGCTGCTGTTGGAGTACATTAGAGTG 300
Db 398 CCTCATGAGGTTGCTGTTCTGTTCTGTCAGAGGCTGCTGTTGGAGTACATTAGAGTG 457
Oy 301 AATGTTACGCGCTTTGTTGTTGAGAACTCACTGTTGATCTTCATCTTCAAGAA 360
Db 458 AACGTCATGCGCTTTGTTGTTGAGAACTCACTGTTGATCTTCATCTTCAAGAA 517
Oy 361 GAGCTGTTGATGA 375
Db 518 GAGCTTGTCTGATGA 532

RESULT 5

US-08-483-376-1
Sequence 1, Application US/08483376

Patent No. 595330
GENERAL INFORMATION:

APPLICANT: Vasil, Vimala

APPLICANT: Clancy, Maureen A.

APPLICANT: Ferl, Robert J.

APPLICANT: Vasil, Indra K.

APPLICANT: Hannah, L. C.

TITLE OF INVENTION: No. 595330e1 Means for Enhancing Gene

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P. C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,376

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/418,540

FILING DATE: 07-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/102,115

FILING DATE: 04-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/830,956

FILING DATE: 05-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/353,854

FILING DATE: 18-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Feider, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 10-94B

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 6386 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Zea mays

STRAIN: Black Sweet

FEATURE:

NAME/KEY: exon

LOCATION: 131..182

FEATURE:

NAME/KEY: exon

LOCATION: 1211..1324

FEATURE:

NAME/KEY: exon

LOCATION: 1828..1948

FEATURE:

NAME/KEY: exon

LOCATION: 2041..2187

FEATURE:

NAME/KEY: exon

LOCATION: 2269..2460

FEATURE:

NAME/KEY: exon

LOCATION: 2605..2728

FEATURE:

NAME/KEY: exon

LOCATION: 2822..3038

FEATURE:

NAME/KEY: exon

LOCATION: 3256..3351

FEATURE:

NAME/KEY: exon

LOCATION: 3447..3620

FEATURE:

NAME/KEY: exon

LOCATION: 3702..3818

FEATURE:

NAME/KEY: exon

LOCATION: 3912..4078

FEATURE:

NAME/KEY: exon

LOCATION: 4158..4381

FEATURE:

NAME/KEY: exon

LOCATION: 4517..4835

FEATURE:

NAME/KEY: exon

LOCATION: 4768..5212

FEATURE:

NAME/KEY: exon

LOCATION: 5372..5510

FEATURE:

NAME/KEY: exon

LOCATION: 5636..5917

US-08-483-376-1

Query Match 6.3%; Score 165.8; DB 2; Length 6386;

Best Local Similarity 55.8%; Pred. No. 1.2e-41; Mismatches 197; Indels 213; Gaps 2;

Matches 517; Conservative 0;

Oy 1294 CAGTGACACATCGCCATGCTTTGGAGAGACAAATATCCAGATTGATCTATTTG 1353

Db 3909 CAGTGACACATCGCCATGCTTTGGAGAGACAAATATCCAGATTGATCTATTTG 3968

Oy 1354 AAGAGCTTGAAGACAAATATCCATTTCTTTGGCAATTTACAGTGAATCTTTTGCATG 1413

Db 3969 GACAAATTTGACAGCCAGTACCACTTCTTTGGCAATTTACAGTGAATCTTTTGCATG 4028

Oy 1414 AACCATACAGATTTCATCATCACAGTACCTTCCAGAGAAATTTGACAGA----- 1461

Db 4029 AACCAACCAATTTTCATCATCACAGTACCTTCCAGAGAAATTTGACAGA----- 4088

Db 240 AGAAGAACCGATCATCTTCTCGATGGCGCTC 272

RESULT 8

US-09-313-294A-1862

Sequence 1862, Application US/09313294A

Patent No. 6476212

GENERAL INFORMATION:

APPLICANT: Ito, Laura Y.

APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

FILE REFERENCE: PL-0017 US

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 1862

LENGTH: 265

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6476212 700551647H1

US-09-313-294A-1862

Query Match 3.3%; Score 85.6; DB 4; Length 265;
Best Local Similarity 60.2%; Pred. No. 4.7e-17;

Matches 159; Conservative 0; Mismatches 104; Indels 1; Gaps 1;

QY 474 GTTCTTAATGTCACCTTTGCGAAATTTGTCATGACAGAGACATGACCTTT 533

Db 1 GTTCTCAACCGAACCTTCTCATCATGTCGCAACAGGAGATGCTTGAAGCCCT 60

QY 534 GCTGAATTCCTCAGATGTCATGTCACAGGAGCAAGATGATGTAATGACAGAT 593

Db 61 GCTGATTTCTCTGTCGTCACCGACAGAGGAGATTTATGATGCT-TATGATAGAT 119

QY 594 TCAGAACTTGATGCTCTTCAACATGTTTGAAGAAAGAGAGATATCTTGATCCCT 653

Db 120 ACAAGCTTGGAGAGGCTTCACTGCTGCTGACCAAGCTAGAGGACCTTGCAAGCT 179

QY 654 ACCTTCGACACCATGTCGCAATTTGAACACCGGTTTCCAGAAATCGTTTGAAG 713

Db 180 CCTGCTGACACCATCTACTCACAATTTGCTTATTAATTTCAAGAGTGGGCTGAGAA 239

QY 714 AGTTGGGCTGACACCGCAAGC 737

Db 240 GTGGGCTGATACAGAGGCAATG 263

RESULT 9

US-09-697-367-1

Sequence 1, Application US/09697367

Patent No. 6323015

GENERAL INFORMATION:

APPLICANT: Orozco Jr., Emil M.

APPLICANT: Calmi, Perry G.

APPLICANT: Meng, Zude

APPLICANT: Tarczynski, Mitchell

TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE

FILE REFERENCE: BB116 US NA

CURRENT FILING DATE: 2000-10-26

PRIOR FILING DATE: 1998-MAY-07

PRIOR APPLICATION NUMBER: PCT/US99/09865

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Microsoft Office 97

SEQ ID NO 1

LENGTH: 2236

TYPE: DNA

ORGANISM: Catalpa speciosa

US-09-697-367-1

Query Match 2.6%; Score 68; DB 4; Length 2236;
Best Local Similarity 48.6%; Pred. No. 1e-10;

Matches 279; Conservative 0; Mismatches 285; Indels 10; Gaps 3;

QY 1612 GAGAGGCGAGGTTGAAGCATTTCCATCTGAGATGCAAGACCTTTTACACCAAGTT 1671

Db 170 GAGACATGATATCTGAGAGTGAAGCAACAGAGTGAAGTCTCCAGAAACACTTATTT 229

QY 1672 GAGAAATGAAGAACTATGCTGCTCAATGACCGCAACGCAATTCCTGTTTCAATG 1721

Db 230 GGGCAGAGGTAATGCTTTCTTTTC---AAATCAAGAGAGCTTATGATTTCTGCACTT 285

QY 1732 CCAAGCTTATGCTGTCAGAACTTAAACCGACTGCTGAGTGGTGGCGCAAGACCA 1791

Db 286 GCCAGGCGAATCCAAAGAAACCTCATCTGTTCAAGCATTTGGGGAATGTCGA 345

QY 1792 AAGTTGCTGATGTTGCTTAACCTGCTAGTGTGATGAGTGAAGAAATCTAA 1851

Db 346 CCACTAAGGAGAGCTGCTTAATCTTAACCTTGAATGAGTAATGATTAATGATGAA 405

QY 1852 GATTGGAAGAGAGGCT---GAAATGAAGAAATGTTGAGCTGATCGAAGTACAA 1908

Db 406 ATGTGGGAACCAATGCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAAT 465

QY 1909 TTGAACGCGCAATTCAGATGATTCATCTCAATGAAGAAATCCGAAATGTTGAATT 1968

Db 466 CTGATGCTGAAGTGGATATCTTAACATCAAGCA---ACATGATGTTCCGAAATTT 522

QY 1969 TACCGATATCTTGCACAGAAAGTCTTGTATGAGCTGATGATGAAGCTTT 2028

Db 523 TACCGTACAGCAAGCAAGAGGCTTTCATTAATCAAGCTTTATTCAGACCTTTT 582

QY 2029 GATTTGACATGTTGAGGAGATGATCTTGCGTTTGCACATTCGAACTGTATGAGCT 2088

Db 583 GGGCTTACTCTCATGAGGCTACAGATATGATGTTGCAATTTGGCAAGAAATGCT 642

QY 2089 GAGCAGCCGAGATTAATGTCATGAGAAATCTGTTTCAACATGATCTTACATGAT 2148

Db 643 GGCCCTGTTGATATACAGAGTTCGCAAGATGCTTCCTGTTGATTCGCCAATACAG 702

QY 2149 GATCAAGCTGCTGACATCTGCTGATTTCTTTG 2182

Db 703 CAGTCAATGCTGATGCTTTTGAAGCTGTTG 736

RESULT 10

US-08-356-354-3

Sequence 3, Application US/08356354

Patent No. 5767365

GENERAL INFORMATION:

APPLICANT: SONNEMALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Ostroienk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,354

FILING DATE: 20-DEC-1994

CLASSIFICATION: 435


```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US PCT/EP93/01605
3      FILING DATE:  22-JUN-1993
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  DE P42 20 758.4
6      FILING DATE:  24-JUN-1992
7      ATTORNEY/AGENT INFORMATION:
8      NAME:  Meliman, Edward A.
9      REGISTRATION NUMBER:  24,735
10     REFERENCE/DOCKET NUMBER:  P/951-105
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE:  (212) 382-0700
13     TELEFAX:  (212) 382-0888
14     TELEX:  236925
15     INFORMATION FOR SEQ ID NO: 3:
16     SEQUENCE CHARACTERISTICS:
17     LENGTH: 3625 base pairs
18     TYPE: nucleic acid
19     STRANDEDNESS: single
20     TOPOLOGY: linear
21     MOLECULE TYPE: cDNA
22     ORIGINAL SOURCE:
23     ORGANISM: Solanum tuberosum
24     FEATURE:
25     NAME/KEY: CDS
26     LOCATION: 121..3282
27     OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
28
29 US-08-356-354-3

```

Query Match	2.3%	Score 59.4	DB 1	Length 3625,
Best Local Similarity	48.0%	Pred. No. 8.5e-08;		
Matches 233;	Conservative	0;	Mismatches 246;	Indels 6;
Gaps				2;
QY	1701	TCACCGCAACAGCCCAATTCGTTCACACATGCGCAAGGCTTGATCGTGCAGAACTTAAC	1766	
DB	1530	TAATTCCAAGGAAGGCTATGATACCTGCACCTGCTAGGCTGTGCCAAGAAAGACCTAC	1583	
QY	1761	CCGACGTCGTGAGTGTGCGGCAAGAACCCAAAGTTGGTGTGGCTTAACCTGTAGT	1822	
DB	1590	TACTTTAGTGAAGCATTTGGTGAATGTCGTCACATTGAGAGAGCTTGCTTAATCTTACTTT	1644	
QY	1821	TGTAGGTGTGATAGCGCAAGAAAGAACTTAAAGATTGGAGAGAGGCTGA--ATGA	1877	
DB	1650	GATATAGGTATCGAGATTAATATGACGAATGTCTAGACCAATTCGACCTTCTCT	1708	
QY	1878	GAAAATGTTGAGCTGATCGACAAGTACAATTGAACGGCCAAATCAATGATATCANTC	1933	
DB	1710	TTCAATCTGAAAATGATAGATAGTATGATCTTTATGCTCAAGTACTTACTTAACA	1766	
QY	1938	TCAAAATGACAAATCCGAAATGTTGAACCTTACCGATACATTTGCGACAAGAAAGGTGC	1997	
DB	1770	CCACAGACAGTCA---GATGTTCCGTATATCTACCGTCTTGCTGCAAAAGACTTAAGGGTGT	1822	
QY	1998	CTTTGTACAGCTTCATTTGATGAAGCCTTTGGAATTGACAGTTGTGAGGCAATACCTTG	205	
DB	1827	TTTATTAATACCTACCTTTATATGAGCCTTTTGAGACTGACTTTGATTGAGGCGACGCTTA	1886	
QY	2058	CGGTTTGCCACATTTGCGACCTGTAAACGTGTGACACACCGAGATTAATGTTCATGGAA	2117	
DB	1887	TGGCTCCCAATGTGTAGCCACAAAAAATGAGAGACCTGTGTATTAACATAGGGTCTTGGA	1946	
QY	2118	ATCTGTGTTCAACATTAATCCTTACATGTGTGATACACCTGTGACATATCTGTGATTT	2177	
DB	1947	CAATGTGCTCTTGAATGATGCCCATGATACAGAGCAATTTGCTGATGCTCTTTTGAAATT	2006	
QY	2178	CTTGTG 2182		
DB	2007	GCTTG 2011		
RESULT 11				
US-08-778-656-3				
Sequence 3, Application US/08778656				

Patent No. 5976869
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
TITLE OF INVENTION: PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,656
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,354
FILING DATE: 20-DEC-1994
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 362-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3625 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: CDS
LOCATION: 121..3282
OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
US-08-778-656-3

	Query Match	2.3%;	Score 59.4;	DB 2;	Length 3625;
	Best Local Similarity	48.08;	Pred. No. 8,56-08;		
	Matches 233;	Conservative 0;	Mismatches 246;	Indels 6;	Gaps 2;
QY	1701	TGACCGAACAAGCCAAATTCGTTCACAAATGCCAAGCGCTGATCGTGCANAACTTAAC	1766		
Db	1530	TAATCCAAAGAAAGCCTATGATACCTCGACATCTGACGCGCTGATCCAAAGAAACCTCAC	1588		
QY	1761	CGAGCTCGTCGAGTGGCGCGGAAACCCAAATTCGTGATGGCTAACCTGAT	1822		
Db	1590	TACTTTAGTGAAAGCAATTTGGTGAATGTCGTCAATGAGAGGCTTGCTAATCTTAATCTT	1648		
QY	1821	TGTAGTGTGATAGCGGAAAGGAATCTAAAGATTTGGAAGAGAGGCTGAA---AATGAA	1877		
Db	1650	GATTAATGGGTATCGAATATATATCGAAGAAATGTCAGACCAATCTGCACCTTCCTT	1709		
QY	1878	GAAATGTTGAGCTGATCGACAAAGTACAACTTGAACGGCCAAATTCAGATGGATATATC	1937		
Db	1710	TTCAATCTTGAATAATGATTAAGTATGATCTTTAATAGTCTCAAGTACCTAATCTTAACA	1768		

QY 1938 TCAATGACAGATCCGAATGTTGAACCTTACCGATACATTTGCCACAGAAAGTGC 1997
DB 1770 CCACAGAGATCA---GATGTTCTCGATATCTACCGTCTGCGAAAGACTAAAGGGTGT 1826
QY 1998 CTTTGACAGCCTGATGTATGAAGCCTTTGATTTGACATTTGAGAGCATGACTTG 2057
DB 1827 TTTTATTATCATGCTTTTATTGAGCCCTTTGAGACTGCTTGAATGAGAGCATTTA 1886
QY 2058 CGGTTTGCAACATTCGCAACCTGTGACGCTGAGACAGCCGAGATTAATTTGCAATGGAA 2117
DB 1887 TGGTCTCCCAATGAGTACCAAAAAATGAGAGACTGTGATATATACATAGGGTCTTGA 1946
QY 2118 ATGCGTTTCAACATGTTGCTTACCATGATGATACAGCTGCTGACATCTGCGATTT 2177
DB 1947 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATGCTTTTGAAGTT 2006
QY 2178 CTTTG 2182
DB 2007 GGTGG 2011

RESULT 12
US-08-356-354-1
Sequence 1, Application US/08356354
Patent No. 5767365
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
NUMBER OF SEQUENCES: 6
PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostroienk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,354
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weilmann, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: E/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3740 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: CDS
LOCATION: 957..3494
OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"

US-08-356-354-1

Query Match 2.3%; Score 59.4; DB 1; Length 3740;
Best Local Similarity 48.0%; Pred. No. 8.7e-08;
Matches 233; Conservative 0; Mismatches 246; Indels 6; Gaps 2;

QY 1701 TGACCGAAGACCAATTCGTTCACATGCAATGCCAGCTTGATGCTGCAAGACTTAAC 1760
DB 1742 TAATCCAGAGAGCTTATGATCTGCACTTGCTAGGCTTGATCCAGAGAACTCAC 1801
QY 1761 CGGACTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCTGATGCTGCAACTCGTAGT 1820
DB 1802 TACTTTAGTGAAGCAATTTGTGATGTCCTTCATTGAGAGAGCTTGCTAATCTTACTTT 1861
QY 1821 TGATGCTGTATGAGCGAAAGAAATCTAAAGATTTGAAAGAGAGCTGAA--ATGAA 1877
DB 1862 GATATAGGGTATTCGATATATATGACGAATGTCTAGCACCNAATTCGCACTTCTTCT 1921
QY 1878 GAAAATGTTGAGCTGATGCAAGTACAACTTGAAAGCGCAATTCAGATGATATCATC 1937
DB 1922 TTCAATCTTGAATAATGATATAGTATGATCTTATATGTCAGATGACTTATCTTAACA 1981
QY 1938 TCAATGACAGATCCGAATGTTGAACCTTACCGATACATTTGCGACAGAAAGTGC 1997
DB 1982 CCACAGAGATCA---GATGTTCTGATATCTACCGTCTTGCCAAAGACTAAGGGTGT 2038
QY 1998 CTTTGACAGCCTGATGTATGAAGCCTTTGATTTGACAGTTGTGAGGCAATGACTTG 2057
DB 2039 TTTTATTATTCAGCTTTTATGAGCCTTTGAGACTGACTTTGATGAGCAGACGCTTA 2098
QY 2058 CGGTTTGCAACATTCGCAACCTGTGACGCTGAGACCGCGAGATTTTGTCCATGGAA 2117
DB 2099 TGGTCTCCCAATGAGTACCAAAAAATGAGAGACTGTGATATATACATGAGGTTCTTGA 2158
QY 2118 ATGCGTTTCAACATGTTGCTTACCATGATGATCAAGCTGCTGACATCTGCGATTT 2177
DB 2159 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATGCTTTTGAAGTT 2218
QY 2178 CTTTG 2182
DB 2219 GGTGG 2223

RESULT 13
US-08-778-656-1
Sequence 1, Application US/08778656
Patent No. 5976869
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
NUMBER OF SEQUENCES: 6
PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostroienk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,656
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,354
FILING DATE: 20-DEC-1994
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993

Db 1386 TTTATTATATCAGCTTTTATTATGACCTTTTGAGTACTTGTGATGGAGACGCTTA 1445
Qy 2058 CGGTTTGCCAACTTCGCACTGTGAACGTGACCAAGCCGAGATTTATTCATGGGAA 2117
Db 1446 TGGCTCCCAATGATGAGCAAAAATGAGAGACCTTTGATATACATAGGTTCTTGA 1505
Qy 2118 ATCTGTTTCAACATTTATCTTACATGATGATCAAGCTCTGACATCTGCTGATTT 2177
Db 1506 CAATGCTCTTGAATGATGCCCATGATCAGCAGGCAATTCGTATGCTTTTGAAGTT 1565
Qy 2178 CTTTG 2182
Db 1566 GGTGG 1570

RESULT 15

US-08-778-656-5

Sequence 5, Application US/08778656

Patent No. 5976869

GENERAL INFORMATION:

APPLICANT: SONNEMALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

NUMBER OF SEQUENCES: 6

ADDRESS: Ostroienk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/778.656

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,354

FILING DATE: 20-DEC-1994

APPLICATION NUMBER: US PCT/EP93/01605

FILING DATE: 22-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P42 20 758.4

FILING DATE: 24-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Meilman, Edward A.

REGISTRATION NUMBER: 24,735

REFERENCE/DOCKET NUMBER: P/951-105

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 382-0700

TELEFAX: (212) 382-0888

TELEX: 236925

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2930 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Solanum tuberosum

FEATURE:

NAME/KEY: CDS

LOCATION: 118..2841

OTHER INFORMATION: /note= "Sucrose-Phosphatase-Synthase"

US-08-778-656-5

Query Match 2.2%, Score 57.8, DB 2, Length 2930,

Best Local Similarity 47.8%; Pred. No. 2,3e-07;
Matches 232; Conservative 0; Mismatches 247; Indels 6; Gaps 2;
Qy 1701 TGACCGCAACAGCAATTCGTTCACAAATGCCAAGCTTGATGTCGAAGAACTTAAC 1760
Db 1089 TAATCCAAAGAGAGCTTATGATTAATCGCACTTGCTAGGCTGATCCAAAGAAAGCTCAC 1148
Qy 1761 CGAAGCTGTGAGTGGTGGCGCAAGAACCCAAAGTTGCGTGAAGTGGCTAATCTGAGT 1820
Db 1149 TACTTTGTAAGACATTTGTGTAATGTCGTCAATGAGAGACCTTGTATCTTACTTT 1208
Qy 1821 TGTAGTGTGATTAAGCGGAAGGAATCTAAAGATTGGAAGAGAGGCTGAA--ATGA 1877
Db 1209 GATTAAGGTAATGAGATTAATATGACGAAGATCTAGCAACCAATTCGACTTCTTCT 1268
Qy 1878 GAAATGTTGAGCTGATGACAAAGTCAAACTTGAAACGCGCAATTCAGATGATATCATC 1937
Db 1269 TTCAATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
Qy 1938 TCAATGAACGAATCCGAATGTTGAATTTGACCTTACCGATTAATTCGACAGAAAGTGC 1997
Db 1329 CCAACAGCAGTCA--GATGTTCTGTATCTACCGTCTTGCCAAAGACTAAAGGCTGT 1385
Qy 1998 CTTTGTACGCTGATTTGATGAGCCTTGATGACAGTTGTGAGGCAATGACTTG 2057
Db 1386 TTTATTATTCAGCTTTTATGAGCCTTTTGACTGACTTGTATGAGGAGCAGCCTTA 1445
Qy 2058 CGGTTTGCCAACTTCGCACTGTGAACGTGACCAAGCCGAGATTTATTCATGGGAA 2117
Db 1446 TGGCTCCCAATGATGAGCAAAAATGAGAGACCTGTTGATATACATAGGTTCTTGA 1505
Qy 2118 ATCTGTTTCAACATTTATCTTACATGATGATCAAGCTCTGACATCTGCTGATTT 2177
Db 1506 CAATGCTCTTGAATGATGCCCATGATCAGCAGGCAATTCGTATGCTTTTGAAGTT 1565
Qy 2178 CTTTG 2182
Db 1566 GGTGG 1570

Search completed: June 5, 2004, 08:36:14
Job time : 223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 00:49:26 ; Search time 1535 Seconds
(without alignments)
7801.459 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625
Sequence: 1 atggctgagcgtgcttcac.....tcaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 295936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2624	100.0	2625	14	US-10-003-405-1 Sequence 1, Appl1
2	1602.6	60.1	5069	13	US-10-424-599-115562 Sequence 115562, Ap
3	1591.6	60.6	2601	13	US-10-425-114-9604 Sequence 9604, Ap
4	1590	60.6	3101	13	US-10-424-599-12144 Sequence 12144, A
5	1570.4	59.8	2772	13	US-10-425-114-6460 Sequence 6460, Ap
6	1570.4	59.8	3360	13	US-10-424-599-11458 Sequence 11458, A
7	1568	59.7	3438	13	US-10-424-599-11459 Sequence 11459, A
8	1566.4	59.7	2772	13	US-10-425-114-9594 Sequence 9594, Ap
9	1499.4	57.1	2906	16	US-10-393-840-139 Sequence 139, Appl
10	1497.8	57.1	2913	16	US-10-393-840-15 Sequence 15, Appl
11	1497.8	57.1	3103	15	US-10-137-036-57 Sequence 57, Appl
12	1486.6	56.6	2427	11	US-09-938-842A-1620 Sequence 1620, Ap
13	1486.6	56.6	2427	11	US-09-938-842A-1620 Sequence 1620, Ap
14	1331	50.7	2279	13	US-10-425-114-9355 Sequence 9355, Ap

15	1305.4	49.7	2714	15	US-10-289-757-133 Sequence 133, Appl
16	1305.4	49.7	2716	15	US-10-289-757-12 Sequence 12, Appl
17	1296.6	49.4	2825	15	US-10-289-757-15 Sequence 15, Appl
18	1295.4	49.3	2754	15	US-10-289-757-13 Sequence 13, Appl
19	1287.6	49.1	2957	15	US-10-289-757-134 Sequence 134, Appl
20	1281	48.8	2908	15	US-10-080-114A-6 Sequence 6, Appl1
21	1277.8	48.7	2986	13	US-10-425-114-2283 Sequence 2283, Ap
22	1277.8	48.7	2986	13	US-10-425-114-26430 Sequence 26430, A
23	1277.8	48.7	2992	13	US-10-425-114-5310 Sequence 5310, Ap
24	1277.8	48.7	3419	13	US-10-425-114-17596 Sequence 17596, A
25	1274.8	48.6	2950	15	US-10-289-757-14 Sequence 14, Appl
26	1260.4	48.0	2671	13	US-10-425-114-31764 Sequence 31764, A
27	1260.4	48.0	2817	13	US-10-425-114-33573 Sequence 33573, A
28	1254	47.8	2746	15	US-10-080-114A-4 Sequence 4, Appl1
29	1233.2	47.0	2127	13	US-10-425-114-10123 Sequence 10123, A
30	1219	46.4	2638	13	US-10-425-114-2574 Sequence 2574, Ap
31	1202.4	45.8	5877	16	US-10-260-238-25 Sequence 25, Appl
32	1190	45.3	2038	13	US-10-425-114-6461 Sequence 6461, Ap
33	1172	44.6	2832	13	US-10-425-114-15113 Sequence 15113, A
34	1150	43.8	3191	13	US-10-424-599-98889 Sequence 98889, A
35	1142.4	43.5	2505	13	US-10-425-114-35609 Sequence 35609, A
36	1115	42.5	2430	9	US-09-938-842A-2415 Sequence 2415, Ap
37	1115	42.5	2430	11	US-09-938-842A-2415 Sequence 2415, Ap
38	1104.8	42.1	2394	15	US-10-217-939-29 Sequence 29, Appl
39	1104.6	42.1	2757	15	US-10-080-114A-11 Sequence 11, Appl
40	1097	41.8	2418	15	US-10-217-939-27 Sequence 27, Appl
41	1096.2	41.8	2737	15	US-10-080-114A-1 Sequence 1, Appl1
42	1041.8	39.7	2210	13	US-10-425-114-8902 Sequence 8902, Ap
43	1031.6	39.3	2307	13	US-10-425-114-35558 Sequence 35558, A
44	982.8	37.4	2035	13	US-10-425-114-5608 Sequence 5608, Ap
45	923.6	35.2	2275	13	US-10-425-114-6036 Sequence 6036, Ap

ALIGNMENTS

RESULT 1
US-10-003-405-1
Sequence 1, Application US/10003405
Publication No. US20020116736A1
GENERAL INFORMATION:
APPLICANT: Ruan, Yong-Ling
APPLICANT: Fuchang, Robert T.
TITLE OF INVENTION: Modification of sucrose synthase gene expression in plant tissue
FILE REFERENCE: GHSUSY WO1
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/251852
PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2625
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1240)..
OTHER INFORMATION: n = any nucleotide (a,g,c,t)
NAME/KEY: CDS
LOCATION: (1)..
OTHER INFORMATION:
US-10-003-405-1
Query Match 100.0%; Score 2624; DB 14; Length 2625;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 ATGGCTGAGCGTGTCTCACTCGGCTCCAGTCTCCGTGAGCGTTGGATGAGACCTT 60
1 ATGGCTGAGCGTGTCTCACTCGGCTCCAGTCTCCGTGAGCGTTGGATGAGACCTT 60

```
OY 61 CTGCTCACAGGACGAGATTTTGCCCTTGCTCTCAAGANTCAGGGCGAAGAAAGCA 120
DB 61 CTGCTCACAGGACGAGATTTTGCCCTTGCTCTCAAGANTCAGGGCGAAGAAAGCA 120
OY 121 ATTGCGAACCACATCAAAATTAATTCTAGAGTTGAGAGCTATCCCGAAGAAACGAAAG 180
DB 121 ATTGCGAACCACATCAAAATTAATTCTAGAGTTGAGAGCTATCCCGAAGAAACGAAAG 180
OY 181 AAGCTCGTAATGATGCAATTTTGAAGATTAAGAGCTAGTCAGAAAGCGATCGTGTG 240
DB 181 AAGCTCGTAATGATGCAATTTTGAAGATTAAGAGCTAGTCAGAAAGCGATCGTGTG 240
OY 241 CCTCATGAGGTGCACTTGCTGTGTCTCAAGGCGCTGTGTGAGAGTCAATAGAGTG 300
DB 241 CCTCATGAGGTGCACTTGCTGTGTCTCAAGGCGCTGTGTGAGAGTCAATAGAGTG 300
OY 301 AATGTTCAAGCCCTGTGTTGAGAGAACCTGTTGCTGATCTCACTTCAAGAA 360
DB 301 AATGTTCAAGCCCTGTGTTGAGAGAACCTGTTGCTGATCTCACTTCAAGAA 360
OY 361 GAGCTTGTTGATGAAAGTTCAAAATGAAACCTTGTTTGAATGGAATTTGAGAGCTTC 420
DB 361 GAGCTTGTTGATGAAAGTTCAAAATGAAACCTTGTTTGAATGGAATTTGAGAGCTTC 420
OY 421 AACTCATATTCCTCCGCGCCAACTCTTTCAAAATCCATTTGATATGATGATGAGTTCCTA 480
DB 421 AACTCATATTCCTCCGCGCCAACTCTTTCAAAATCCATTTGATATGATGATGAGTTCCTA 480
OY 481 AATGTCACCTTTGGGCAAAATTTGTCATGACAGAGAGACATCCCTTTGTCGAA 540
DB 481 AATGTCACCTTTGGGCAAAATTTGTCATGACAGAGAGACATCCCTTTGTCGAA 540
OY 541 TTCTCAGAGTCAATGTCACAAAGGCAAGAACATGATGTTGAATGACAGAAATTCAGAAC 600
DB 541 TTCTCAGAGTCAATGTCACAAAGGCAAGAACATGATGTTGAATGACAGAAATTCAGAAC 600
OY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGATCTTGTATCCCTACCTCT 660
DB 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGATCTTGTATCCCTACCTCT 660
OY 661 GAGACACCATGTCGCGAATTCGAACAACCGGTTCCAGAAATCGGTTTGAAAGAGTTGG 720
DB 661 GAGACACCATGTCGCGAATTCGAACAACCGGTTCCAGAAATCGGTTTGAAAGAGTTGG 720
OY 721 GATGACACCGGAGAACGCGTGTCTGAGATGATCCAACTCTTTTGAATCTTCTGAGGCA 780
DB 721 GATGACACCGGAGAACGCGTGTCTGAGATGATCCAACTCTTTTGAATCTTCTGAGGCA 780
OY 781 ACTGATCCTTGCAACCTTGAGAAAGTTCTTGAGAAATCCCAATGATGATGATG 840
DB 781 ACTGATCCTTGCAACCTTGAGAAAGTTCTTGAGAAATCCCAATGATGATGATG 840
OY 841 ATTCTCATCTCCCAAGGATACCTTGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 900
DB 841 ATTCTCATCTCCCAAGGATACCTTGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 900
OY 901 GGGCAGGTGTTTACATCTTGATCAAGTCCGAGCTTGAGAAATGAGATGCTCTCCGT 960
DB 901 GGGCAGGTGTTTACATCTTGATCAAGTCCGAGCTTGAGAAATGAGATGCTCTCCGT 960
OY 961 ATAAAGCAACAAGACTCAACATCAACCTCGAATCTCATTAATTAAGACTTCTTCT 1020
DB 961 ATAAAGCAACAAGACTCAACATCAACCTCGAATCTCATTAATTAAGACTTCTTCT 1020
OY 1021 GATGCTGTGCGAACAACATGCGGTCAAGACTTGAGAAAGTATACGGAACGAGACCTCG 1080
DB 1021 GATGCTGTGCGAACAACATGCGGTCAAGACTTGAGAAAGTATACGGAACGAGACCTCG 1080
OY 1081 GATATCTTTCGAGTACCTTTCAGAACAGAAAGGAAATTTGTCAGAAATGATCTCAGAA 1140
DB 1081 GATATCTTTCGAGTACCTTTCAGAACAGAAAGGAAATTTGTCAGAAATGATCTCAGAA 1140
OY 1141 TTTGAAAAAGTCTGCGCATATCTTGAAAACTTACACAGAGATGTTGCTCATGAAATCTCC 1200
DB 1141 TTTGAAAAAGTCTGCGCATATCTTGAAAACTTACACAGAGATGTTGCTCATGAAATCTCC 1200
OY 1201 AAAGGTTGACGGGACGCGCATCTGATCATCGGAAACNACAGGCAATGATGCTC 1260
DB 1201 AAAGGTTGACGGGACGCGCATCTGATCATCGGAAACNACAGGCAATGATGCTC 1260
OY 1261 GCCTCTGTGTCGACATAAATTTAGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
DB 1261 GCCTCTGTGTCGACATAAATTTAGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
OY 1321 AAGCAAAATATTCAGATTCATATCTATTGGAAGAACTTGAAGCAATTAACATTTG 1380
DB 1321 AAGCAAAATATTCAGATTCATATCTATTGGAAGAACTTGAAGCAATTAACATTTG 1380
OY 1381 TCTTGCCAAATTAAGCGATCTTTTGCAATGAAACCATTAAGATTCATCATACCAAGT 1440
DB 1381 TCTTGCCAAATTAAGCGATCTTTTGCAATGAAACCATTAAGATTCATCATACCAAGT 1440
OY 1441 ACTTTCAGAAATTTGACAGAGAGACACATGTTGTCATATAGAGGACACATGCT 1500
DB 1441 ACTTTCAGAAATTTGACAGAGAGACACATGTTGTCATATAGAGGACACATGCT 1500
OY 1501 TTCACTCTTCTGCTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATG 1560
DB 1501 TTCACTCTTCTGCTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATG 1560
OY 1561 AACTGTTTCCCTGGAGTGAATGAGATATATCTTCCCTTACCGAAGAAAGCGG 1620
DB 1561 AACTGTTTCCCTGGAGTGAATGAGATATATCTTCCCTTACCGAAGAAAGCGG 1620
OY 1621 AGGTTGAAGATTTTCATCTGATGAGAGACCTTCTTTACACAAAGTTGAGATGAA 1680
DB 1621 AGGTTGAAGATTTTCATCTGATGAGAGACCTTCTTTACACAAAGTTGAGATGAA 1680
OY 1681 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
OY 1741 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
DB 1741 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
OY 1801 GAGTTGCTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 1801 GAGTTGCTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATG 1860
OY 1861 GAGGAGCTGAAATGAGAAATGTTGAGCTGATGACAAATGATGAAACGCGCAA 1920
DB 1861 GAGGAGCTGAAATGAGAAATGTTGAGCTGATGACAAATGATGAAACGCGCAA 1920
OY 1921 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
DB 1921 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
OY 1981 TGCGACAGAAAGTGCCTTGTACAGCTGATGATGATGATGATGATGATGATGATG 2040
DB 1981 TGCGACAGAAAGTGCCTTGTACAGCTGATGATGATGATGATGATGATGATGATG 2040
OY 2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
DB 2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
OY 2101 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
DB 2101 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
OY 2161 GACATATCTGCTGATTTCTTTGAAAAAGTGAAGAAATCATATCTGAGGATTAAGATC 2220
DB 2161 GACATATCTGCTGATTTCTTTGAAAAAGTGAAGAAATCATATCTGAGGATTAAGATC 2220
OY 2221 TCCCAAGAGGCTTGAAACGAATGAGAGAGATATATCATGAGAGATTTACTCGAAGAG 2280
```


Db 2221 TCCCAAGGAGCTGAACGATAGAGAGATATACATGGAATTTACTCGGAGAGA 2280
Qy 2281 CTATTGACCTGACAGAGGTATATGATTTCTGGAAGCATGTTTCAACCTTGAAGCCGT 2340
Db 2281 CTATTGACCTGACAGAGGTATATGATTTCTGGAAGCATGTTTCAACCTTGAAGCCGT 2340
Qy 2341 GAGAGTGTCTGATCTGAGATGTTTATGCTCTTAAGTACCCGTAGCTGCTGATCA 2400
Db 2341 GAGAGTGTCTGATCTGAGATGTTTATGCTCTTAAGTACCCGTAGCTGCTGATCA 2400
Qy 2401 GTTCATTTGACAGAGAGATTAATGACCTGTTAAATACATGAGCGGTTTCTTGG 2460
Db 2401 GTTCATTTGACAGAGAGATTAATGACCTGTTAAATACATGAGCGGTTTCTTGG 2460
Qy 2461 AGAATATATTTCTGTTTGTATTTCAATGAGAGAGCTCTTGTATTTCACTTGTCT 2520
Db 2461 AGAATATATTTCTGTTTGTATTTCAATGAGAGAGCTCTTGTATTTCACTTGTCT 2520
Qy 2521 TTTCCTTTTCTTTTGGCGGATGTTTGAACATGAGGAGTGGCGCCGTCAATTCC 2580
Db 2521 TTTCCTTTTCTTTTGGCGGATGTTTGAACATGAGGAGTGGCGCCGTCAATTCC 2580
Qy 2581 AGTTAAATATGATGATCTTTGTTTCAAAAAA 2625
Db 2581 AGTTAAATATGATGATCTTTGTTTCAAAAAA 2625

RESULT 2

US-10-424-599-115562

Sequence 115562, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OR INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 115562

LENGTH: 5069

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_75363C.1

US-10-424-599-115562

Query Match 61.1%; Score 1602.6; DB 13; Length 5069;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 1918; Conservative 0; Mismatches 505; Indels 3; Gaps 1;

Qy 1 ATGGGTGAGCGTCTCTCACTGCGCTCCACAGTCTCCGTGAGCGTTTGATAGACCTT 60
Db 500 ATGGGCACCGATCGTTTACCCGGGTTTCAAGTCTCCGTGAGCGTTTGATAGACCTT 559
Qy 61 CTGCTCAGAGAGAGATTTTGGCTTGTCTCAAGATGAGGCAAGAAAGAA 120
Db 560 ACTGCAACAGAGATGAATTTTGGCTTGTCTCAAGATGAGGCAAGGCAAGG 619
Qy 121 ATTCTGCAACACCATCAATATTTCTAGAGTTTGAAGTATCTCTGAAGAAAG 180
Db 620 ATCTGCAACACCAACAGGTCTATGCTGAGTTTGAAGAAATCCCTGAAGAAAG 679
Qy 181 AAGCTCGTAAATGAGCAATTTTGAAGTGAAGGCTAGTCAGAGAGCAATGCTGT 240
Db 680 AAGCTCGTAAATGAGCAATTTTGAAGTGAAGGCTAGTCAGAGAGCAATGCTGT 739
Qy 241 CCTCATGGGTTGCACTTGTCTGCTCAAGGCGCTGAGTTTGGAGATACATTAGAGTG 300
Db 241 CCTCATGGGTTGCACTTGTCTGCTCAAGGCGCTGAGTTTGGAGATACATTAGAGTG 300

Db 740 CCACATGAGGTTGCTGTGCTGTCTGCTCAAGACCTGTGTGTGGAGTACCTGAGAGTG 799
Qy 301 AATGTACAGCCCTTGTGTTGAGAACTCACTGTGCTGAGTATCTTCAAGGAA 360
Db 800 AATGTACAGCCCTTGTGTTGAGAACTGAGAACTGCTGAGTATCTTCAAGGAA 859
Qy 361 GAGCTTGTGAGAAAGTTCAATGAAATCTTGTGAAATTTGATTTGAGCCCTTC 420
Db 860 GAGCTTGTGAGAAAGTTCAATGAAATCTTGTGAGTATCTTCAAGGAAATTC 919
Qy 421 AACTATCATTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 480
Db 920 AATGAGGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTC 979
Qy 481 AATGAGGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 540
Db 980 AATGAGGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTC 1039
Qy 541 TTCTCAGAGTCAATGTCACAAAGGAGCAAGATGTTGAATGACAGAAATTCAGAAC 600
Db 1040 TTCTCAGAGTCAATGTCACAAAGGAGCAAGATGTTGAATGACAGAAATTCAGAAC 1099
Qy 601 TTGAATGCTCTTCAATGTTTGAAGAGAGAGATCTTGTGTAACCTTACTCTT 660
Db 1100 CCAGATGCACTCAAACTGTTCTGAGAAAGCTGAGAGATCTGAGGACAGTGCCTCT 1159
Qy 661 GAGACACATGTCGCAATTTGAAACACCGTTTCAAGAAATCGTTTGAAGAGTTGG 720
Db 1160 GAAACTCTTCACTCAAGATTTGAGCAAGTTTCAAGATTTGTTGAGAGAGGAGTG 1219
Qy 721 GATGACACGAGAGAGCGGCTCTGAGATGATCAACTCTTGTGATCTTCTTGAAGCA 780
Db 1220 GATGACACGAGAGAGCGGCTCTGAGATGATCAACTCTTGTGATCTTCTTGAAGCA 1279
Qy 781 ACTGATCTTCACTGAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 840
Db 1280 CTTGACCGGTCACCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 1339
Qy 841 ATTCTCACTCTTCACTGAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 900
Db 1340 ATTCTCTTCCCACTGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAG 1399
Qy 901 GAGGAGGTTGTTTCACTGATGATCAAGTCCGAGCTTGAAGAGTCTTGAAGAGTCTTGA 960
Db 1400 GAGGAGGTTGTTTCACTGATGATCAAGTCCGAGCTTGAAGAGTCTTGAAGAGTCTTGA 1459
Qy 961 ATTAAGCAACAGAGCTCAACATCAACCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 1020
Db 1460 ATTAAGCAACAGAGCTCAACATCAACCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 1519
Qy 1021 GATGCTGTGAGAAACATGCGGTCAACGATTTGAGAAAGTATACGAAACAGAGACTG 1080
Db 1520 GATGCTGTGAGAAACATGCGGTCAACGATTTGAGAAAGTATACGAAACAGAGACTG 1579
Qy 1081 GATGCTGTGAGAAACATGCGGTCAACGATTTGAGAAAGTATACGAAACAGAGACTG 1140
Db 1580 GATGCTGTGAGAAACATGCGGTCAACGATTTGAGAAAGTATACGAAACAGAGACTG 1639
Qy 1141 TTTGAAAAAGTCTGAGCAATCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAG 1200
Db 1640 TTTGAAAAAGTCTGAGCAATCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGAAGAGTCTTGA 1696
Qy 1201 AAAGAGTTGACAGGCAAGCTGATCTGATCACTGAGAAACAGAGAGGCAATTCCT 1260
Db 1697 AAAGAGTTGACAGGCAAGCTGATCTGATCACTGAGAAACAGAGAGGCAATTCCT 1756
Qy 1261 GCTCTCTGCTGCAATTAATAGTGTCAACAGTGCACATGCGCCATGCTTGGAG 1320
Db 1757 GCTCTCTGCTGCAATTAATAGTGTCAACAGTGCACATGCGCCATGCTTGGAG 1816
Qy 1321 AAGACAAATATTCAGATTCAGATATCTATTTGAAGAGAGCTTGAAGCAATTCATTC 1380
Db 1817 AAGACAAATATTCAGATTCAGATATCTATTTGAAGAGAGAGCTTGAAGCAATTCATTC 1876

Oy	1381	TCCTGCCAATTTCACGCTGATCTTTTGCATATGAAACAATACAGATTTCATCATACCAAGT	1440
Db	1877	TCCTGCCAATTTCACGCTGATCTTATTGCGATGAAACAACAAGATTTCATTAACCAAGT	1936
Oy	1441	ACTTTCGAGAAATTCGAGAAACGAGACACTGTTGTCATAATCGAGGCCACTGCT	1500
Db	1937	ACCTTCGAGAGATTGCTGAAGACGAGACACTGTTGACAGTACGAATCTCCACAGCC	1996
Oy	1501	TTCACTCTTCCTGGCTCTCAACCGTGTGTGTAATGCTATCCGATGTGATCCCAATTC	1560
Db	1997	TTCACTCTTCCTGGACTCTACCGCTGTGATGTGATATGATGCTTTGATCCAAATTC	2056
Oy	1561	AACATTGTTCCCGCGTGTCTGATATGAGATAATCTTCCCTTACACCGAAGAGAGCGG	1620
Db	2057	AACATTGCTCCCGCGAGCTGATCAAAACATTTACTTCCCCACACTGAACAACAGCCGT	2116
Oy	1621	AGTTGAAGCAATTCATCTGAGATGAGAACCTTCTTTACACCAAGTTGAGATGAA	1680
Db	2117	AGTTGAATCATCTTCCACCTCGAATTCAGAAACCTCTTTACAGCTCAGTGAAGATGAA	2176
Oy	1681	GAAACCTTATGTGTGCTCAATGACCGCAACAAGCCAAATTCGTGTCACATGCCAAGGTT	1740
Db	2177	GAAACCAATGTGTGTCTGAAGACCGCAGCAAGCCAAATTAATCTTCACATGCAAGGTTG	2236
Oy	1741	GATCGTGTCAAGAACTTAACCGGACTCGTCAGTGTGTCCGGCAAGAACCCAAAGTGTGCT	1800
Db	2237	GATCGAATGAAAGAACATCAACAGACTTGTGAGTGTGTACGGTAAAGAACCGAAGCTGAG	2296
Oy	1801	GAGTTGCTAACTCTGTAGTGTGTAGTGATGAGCAAGAAATCTTAAGATTTGGAA	1860
Db	2297	GAGCTGTGAAACCTTGTGTGTGCTGGAACAGAGAGAGAGTCAAGACTTGGA	2356
Oy	1861	GAGAAAGCTGAATGAAGAAATGTTTGAAGCTGATCGCAAGTCAACTTGAACGGCCAA	1920
Db	2357	GAAAGAGCCGAGATGAAGAAATGTGACGGCTGATCGAGACTTCAGATTGAACGGCCAA	2416
Oy	1921	TTTCAGATGATATCATCTCAATATGAACAGAAATCCGAATGTTGAACTTTACGATACTT	1980
Db	2417	TTTCAGATGATTTTCAATGCGAATGAACGATGTGAGGAATGGAAGAGCTCTACCGGTATC	2476
Oy	1981	TGCGACACGAAAGGTGCTTTGTATACGCTTGCAATTTATGAAGCTTTGATTGATCAAGTT	2040
Db	2477	TGCGACACAGAGGGGTGCTTTGTGTGACGCTCTGTATCGAGGCTTTGTGATTGACAGTG	2536
Oy	2041	GTGAGAGCAATGACTTGAGGTTCGCAACTTCGCAACTGTAAACGGTGAACGAGCCGAG	2100
Db	2537	GTGAGGCAATGACTTGCGGCTTCGCAACTTCGCAATGCAATGGAATGGGTGCTGCTGAG	2596
Oy	2101	ATTATTTGTCAATGGGAAATCTGTTTCAAAATGATCTTTACCAATGATCAATGATCAAGTCT	2160
Db	2597	ATTATTTGTGACCGGAAATCTGCGCTTCACATTTGACCTTTACATGTGACCGGTGTCT	2656
Oy	2161	GACATATCTGCTGATTTCTTTGAAAGTGTATGAAGATTCATCTCACTGAGGATTAAGATC	2220
Db	2657	GATCTCCCTTGTGTGACTTTTGAAGAAATGCAAGCTTGACCCAACTCACTGGGAACAAATC	2716
Oy	2221	TCCCAAGAGGCTTGAAGCAATGAGAGAAAGTATACATGGAAGATTTATCTGGAGAGAA	2280
Db	2717	TCAAAAGCTTGATCTTCCAGCGTATGGAAGAAATGACATATGCAAAATTTACTTCAAGG	2776
Oy	2281	CTATTGACCTTGACAGAGTATGGAATTCGGAAGCAATGTTTCCAACTTTGAACGGCGT	2340
Db	2777	CTTCTCACTCTCAACGGGTGTCTATGGCTCTTGAAGACATGTGTAACTTTGAACGGCGT	2836
Oy	2341	GAGAGTGTGCTTACCTTTGAAGATTTTATGCTCTTAATGATACGTAAGCTGTGAAATCA	2400
Db	2837	GAGAGCGCGCCCTATCTCGAGATGTTTATGCTCTCAAGTACCGCAAAATGCTGAGTCT	2896
Oy	2401	GTTCCATTTGAGAGAGAGATTAATGAA	2426
Db	2897	GTGCCCCCTTGCTGCTGAATTAACCTGA	2922

Query	Match	60.6%	Score 1591.6	DB 13	Length 2601
Best Local Similarity	78.7%	Pred. No. 0	Mismatches 515	Indels 3	Gaps 1
Matches 1913	Conservative	0			
US-10-425-114-9604					
Sequence 9604	Application US/10425114				
Publication No.	US2004003488B1				
GENERAL INFORMATION:					
APPLICANT:	Liu, Jindong				
APPLICANT:	Zhou, Yihua				
APPLICANT:	Kovalic, David K.				
APPLICANT:	Screen, Steven E				
APPLICANT:	Tabaska, Jack E				
APPLICANT:	Cao, Yongwei				
TITLE OF INVENTION:	Nucleic Acid Molecules and Other Molecules Associated With				
TITLE OF INVENTION:	Plants and Uses Thereof for Plant Improvement				
FILE REFERENCE:	38-21(5313)B				
CURRENT APPLICATION NUMBER:	US/10/425,114				
CURRENT FILING DATE:	2003-04-28				
NUMBER OF SEQ ID NOS:	73128				
SEQ ID NO 9604					
LENGTH:	2601				
TYPE:	DNA				
ORGANISM:	Glycine max				
FEATURE:					
OTHER INFORMATION:	Clone ID: 700852943_FLI				
US-10-425-114-9604					
Query Match	60.6%	Score 1591.6	DB 13	Length 2601	
1	ATGGCTGAGCGTGTCTCTCACTCGCGCTCCACAGTCTCCGTGAGCGTTTGGATGAGACCTT	60			
40	ATGGCACTGATCGTTTGAACCCGGGTTTCAAGTCTCCGTGAGCGTTTGGATGAGACCTT	99			
61	CTTGTCTCAGAGAAAGAGATTTTGGCCTTCTCTCAAGATGAGGCGAAAGAAAGAA	120			
100	ACTGCGCAACAGAAAGAAATTTTGGCCTTCTCTCAAGATGAGGCGAAAGAAAG	159			
121	ATTCTGCAACACCATCAATATTATCTAGATTGAAGCTATCCCTGAAAGAAAGAAAG	180			
160	ATCTGCAACACCATCAAGTCTCTGATTTGAAGAAATCCCTGAAAGAAAGAAAG	219			
181	AAGCTCGTAAATGATGATTTTGAAGTATGAAAGGCTAGTCAGAAAGCAATCTGTTT	240			
220	AAGCTTACTGATGATGATTTTGAAGTATGAAAGGCTAGTCAGAAAGCAATCTGTTT	279			
241	CCTCTGATGATGATGATTTTGAAGTATGAAAGGCTAGTCAGAAAGCAATCTGTTT	300			
280	CCACCAATGATGATGATTTTGAAGTATGAAAGGCTAGTCAGAAAGCAATCTGTTT	339			
301	AATGCTCAGCGCTTGTGTTGAGGAACTCATCTGTGCTGATATCTCCATTCAGAA	360			
340	AATGCTCAGCGCTTGTGTTGAGGAACTCATCTGTGCTGATATCTCCATTCAGAA	399			
361	GAGCTTGTGATGAGAACTTCAATGAGAACTTTGTTTGAATTTGAAATTTGAGCCCTT	420			
400	GAACTTGTGATGAGAACTTCAATGAGAACTTTGTTTGAATTTGAAATTTGAGCCCTT	459			
421	AATCTATATTTCCCGCGCGCAATCTTCAAAATTCATTTGTAATGATGAGATTTCTTA	480			
460	AATCTATATTTCCCGCGCGCAATCTTCAAAATTCATTTGTAATGATGAGATTTCTTA	519			
481	AATCTATATTTCCCGCGCGCAATTTGTTGATGAGAAAGGAGATGAGATTTGAGCCCTT	540			
520	AATCTATATTTCCCGCGCGCAATTTGTTGATGAGAAAGGAGATGAGATTTGAGCCCTT	579			
541	TTCTCTCAGAGCTTCAAGGCGCAATTTGTTGATGAGAAAGGAGATGAGATTTGAGCCCTT	600			
580	TTCTCTCAGAGCTTCAAGGCGCAATTTGTTGATGAGAAAGGAGATGAGATTTGAGCCCTT	639			
601	TTGATATGCTTCTTCAAGCTTTTGAAGAAAGGAGATGAGATTTGAGCCCTT	660			
640	TGATATGCTTCTTCAAGCTTTTGAAGAAAGGAGATGAGATTTGAGCCCTT	699			

```

OY 661 GAGCACCATGTCGGAATTCGAACACCGGTTCCAGAAATCGGTTGAAAAGAGTTGG 720
DB 700 GAACTCCCTACTGAAATTTGAGACAAAGTTCCAGAGATTGGTTTGAGAGAGGGTGG 759
OY 721 GGTGACACCGAGAAACGCGTGTGAGATGATCCAACTCTTTTGATCTTTGAGGCA 780
DB 760 GGTGACAAACGAGCGTGTCTTGATGATCAATCTCTTGATCTTTGAGAGCC 819
OY 781 ACTGATCTTGCACTCTTGAGAAAGTTCTTGAGAAATCCCAAGGTTCAAGTTTGG 840
DB 820 CCGACCTTGCACTCTTGAGAAAGTTCTTGAGAAATCCCAAGGTTCAAGTTTGG 879
OY 841 ATTCTCATCTCCCAAGGATTAATCTGCTCAAGCAATGTTTGAGGATATCCGACCGGT 900
DB 880 ATTCTCTTCCCAAGGATTAATCTGCTCAAGCAATGTTTGAGGATATCCGACCGGT 939
OY 901 GGGCAGGTTGTTTACATCTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTCTCGT 960
DB 940 GGGCAGGTTGTTTACATCTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTCTCGT 999
OY 961 ATTAAGCAACAGAGCTCAATCACTACCCCTGAAATCTCTTAATTAAGTACTTCTCT 1020
DB 1000 ATTAAGCAACAGAGCTCAATCACTACCCCTGAAATCTCTTAATTAAGTACTTCTCT 1059
OY 1021 GATGCTGTGGAACAAACATGCGGTCAAGCACTTGAGAAAGTATAGAAACAGAGCACTCG 1080
DB 1060 GATGCTGTGGAACAAACATGCGGTCAAGCACTTGAGAAAGTATAGAAACAGAGCACTCG 1119
OY 1081 GATATCTCTGAGTACCTCTTGAGAAACAGAAAGGATTTGTTGAGAAATGATCTCAAGA 1140
DB 1120 CACATCTCTGAGTACCTCTTGAGAAACAGAAAGGATTTGTTGAGAAATGATCTCAAGA 1179
OY 1141 TTTGAAAAAGTCTGCGCACTCTTGAGAAACCTACAGAGATGTTGCTCAAGAAATCTCC 1200
DB 1180 TTTG----AAGTCTGCGCACTCTTGAGAAACCTACAGAGATGTTGCTCAAGAAATCTCC 1236
OY 1201 AAAGAGTTGCAACGCGCAACGCACTGATCATTCGAAACCAAGGAGCGCAATTCGTC 1260
DB 1237 AAAGAGTTGCAACGCGCAACGCACTGATCATTCGAAACCAAGGAGCGCAATTCGTC 1296
OY 1261 GCTCTCTGCTGCACTAATTAAGTGTCAACAGTGCACCACTGCTTGGAG 1320
DB 1297 GCTCTCTGCTGCACTAATTAAGTGTCAACAGTGCACCACTGCTTGGAG 1356
OY 1321 AAGACAAATATCCGATTCAGATATCTATTGGAAGAACTTGAAAGCAATTCATTC 1380
DB 1357 AAGACAAATATCCGATTCAGATATCTATTGGAAGAACTTGAAAGCAATTCATTC 1416
OY 1381 TCTTGCAATTTTACGCTGATCTTTTGCATGAAACCAATGATTTGATCATCACTGAGT 1440
DB 1417 TCTTGCAATTTTACGCTGATCTTTTGCATGAAACCAATGATTTGATCATCACTGAGT 1476
OY 1441 ACTTTCAGAAATTTGAGAAAGCAAGCACTGTTGTTCAATACAGAGCCACTGCT 1500
DB 1477 ACTTTCAGAGATTTGAGAAAGCAAGCACTGTTGTTCAATACAGAGCTCTACACAGCC 1536
OY 1501 TTTACTCTTCTGCTGCTACCGTGTGTAATGATGATGATGTTGATCTCAATTC 1560
DB 1537 TTTACTCTTCTGCTGCTACCGTGTGTAATGATGATGATGTTGATCTCAATTC 1596
OY 1561 AAACATGTTTCCCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1597 AAACATGTTTCCCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1656
OY 1621 AGTTTGAAGCAATTTTCATCTGAGATGAGAAAGCTTTCTTTTACCAAAAGTTGAGATGA 1680
DB 1657 AGTTTGAAGCTCTCTTCAACCCGAAATTCGAAGAACTTTTACAGCTCTGTGAGAAATGA 1716
OY 1681 GAACTTATGTTGCTCAATGACGCAACAGCAATTTGTTTCAATGCAATGCAAGGCTT 1740
DB 1717 GAACTTATGTTGCTCAATGACGCAACAGCAATTTTCAATGCAATGCAAGGCTT 1776

```

```

OY 1741 GATGCTGTCAAGAACTTAACCGGACTGCTGAGTGTGTCGGCAAGAACCCAAAGTTGCT 1800
DB 1777 GACCTGTGAAGAACATCAACAGACTGCTGAGTGTGTAAGGTAAGAACCGAAGCTGAG 1836
OY 1801 GATTTGGCTAACCTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1837 GATTTGGTGAACCTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1896
OY 1861 GAGAAAGCTGAATGAGAAATGTTGAGCTGATGCAAGTCAACTGAACGCGCA 1920
DB 1897 GAGAAAGCTGAATGAGAAATGTTGAGCTGATGCAAGTCAACTGAACGCGCA 1956
OY 1921 TTCAATGATATCATCTCAATGAACGAATCCGAATGTTGAATTTACCGATACAT 1980
DB 1957 TTCAATGATATTTATCTCATGATCAATGAACCGTGTGAGAAACGAGCTGTGATC 2016
OY 1981 TCGCACAGAAAGTCTCTTGTGATACGCTGCAATTTGATGAGCTTTGATGACAT 2040
DB 2017 TCGCACACCAAGGAGCTTTGTGATGAGCTGCAATTTGATGAGCTTTGATGACAT 2076
OY 2041 GTGAGGCAATGATGCTGCGGTTTGCAACATTCGCAACCTGTAACGCTGACAGCGAG 2100
DB 2077 GTGAGGCAATGATGCTGCGGTTTGCAACATTCGCAACCTGTAACGCTGACAGCGAG 2136
OY 2101 ATTAATGCTCATGAGAAATCTGTTTCAACATGATCTTACATGATGATCAAGTCT 2160
DB 2137 ATTAATGCTCATGAGAAATCTGTTTCAACATGATCTTACATGATGATCAAGTCT 2196
OY 2161 GACATCTCTGCTGATTTCTTTGAAAGTGTAGAAAGTCAATCTCACTGGATTAAGTC 2220
DB 2197 GATCTCTTGTGATCTTCTTTGAGAAAGTCAAGCTGACCCCACTGAGAAACATC 2256
OY 2221 TCCCAAGGCTTGAACGAATGAGGAGAAATGATGATGATGATGATGATGATGATGATGAT 2280
DB 2257 TCCCAAGGCTTGAACGAATGAGGAGAAATGATGATGATGATGATGATGATGATGATGAT 2316
OY 2281 CTATGACCTCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2317 CTATGACCTCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2376
OY 2341 GAGAGCTGCTGATACCTTGAGATGTTTATGCTTTAAGTACCTGAGCTGATCA 2400
DB 2377 GAGAGCTGCTGATACCTTGAGATGTTTATGCTTTAAGTACCTGAGCTGATCA 2436
OY 2401 GTTCATTTGCAAGAGGATTAATTGAACCTG 2431
DB 2437 GTTCATTTGCTGTTGATTAATGAGATG 2467

```

RESULT 4

```

US-10-424-599-12144
; Sequence 12144, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12144
; LENGTH: 3101
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110974C.1
US-10-424-599-12144
Query Match 60.6%; Score 1590; DB 13; Length 3101;

```

Best Local Similarity 78.7%; Pred. No. 0; Matches 1912; Conservative 0; Mismatches 516; Indels 3; Gaps 1;

1 ATGGCTGACGCTGCTCACTGCGCTCCAGCTCCCGTAGCGTTGATGAGACCTT 60
131 ATGGCCACTGATCTGTTGACCGGGTTCAAGCTCTCCGTGAGAGCTTGAGAACCTTC 190
61 CTGTGCTACAGGAACGAGATTTTGGCCCTTGTCTCAGAGATCGAGGGCAAGAAAGGA 120
191 ACTGCCAACAGGAATTTTGGCCCTTGTCTCAGAGATCGAAGCTTAAGGGCAAGGGG 250
121 ATTCTGCAACACATCAATTAATTCTAGAGTTTGAAGCTATCCCTGAAGAGAACAGAAAG 180
251 ATCTGCAACACACAGAGTCAATGCTGAGTTTGAAGAAATCCCTGAGAGAACAGGAG 310
181 AAGCTGCTAATGAGTCAATTTTGTGAAATGAAAGAGTCAAGAGAGCGATCGTGTG 240
311 AAGCTTACTGATGAGTCCCTTTGGAGAGTCTTGAGATCTACAGAGAACCTAAGTTTGG 370
241 CCTCATGAGGTTCACATTTGCTGTTGTCACAGGCGCTGAGTTTGGAGTACATTAGAGTG 300
371 CCACCATGGGTGCTGCTGGCTGTTGTCACAGGCGCTGAGTGTGGAGTACCTGAAAGTG 430
301 AATGTTACGCGCTTGTGTTGTTGAGAACTCACTGTTGCTGAGTATCTCACTTCAAGGA 360
431 AATGTCACGCTCTGTTGTTGTTGAGAGTTCACACCTGCTGAGTACCTTCACTTCAAGGA 490
361 GAGCTTGTGATGAGAGTTTCAAAATGAAACCTTTGTTTGAATTTGAGATTTGAGCCCTTC 420
491 GAACTTGTGATGAGAGTTTCTAATGGAACCTTTGTTGAGTTGAGATTTGAGAACATTC 550
421 AACTCATGATTTCCCGCCCAACTCTTTCAAAATCATTTGTAATGAGTGTGAGATTCCTA 480
551 AATGACGCTTCCCTCGCCCAACTCTTAACAAGTCAATTTGAAATGAGTGTGACATTCCTC 610
481 AATGTCACCTTTGGGCAAAATTTGTTCAATGACAGAGAGAGACATCCCTTGTGCA 540
611 AACGCCACCTTTGCGCAAACTTTCCACAGCAGAGAGAGCTTGACACCTTTTGGAA 670
541 TTCTCAGAGTCCATTTGTCACAGAGGCAAGAACATGATTTGAATGACAGAAATCAGAAC 600
671 TTCTCAGAGCTTCAACGCTGCAAGGAAAGACTTGAATTTGAATGACAGAAATTCAGAAC 730
601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGATCTTGTATCCCTACCTCT 660
731 CCAATGACATCCAAATGTTCTGAGAAAGCTGAGATCTTGTGGCAAGTGTCTCT 790
661 GAGACACCATGTCGGAATTTGGAACACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG 720
791 GAAACTCCCTACTCAAGATTTGAGCAAAATTCAGAGATTTGTTGGAAGAGGGGTGG 850
721 GGTGACACCGGAGACGCGTGTGAGATTTCCAACTCTTTTGGATCTTCTTGAAGCA 780
851 GGTGACACCGAGAGCGTGTGTTGATGATCAATTCATCTCTTGTGATCTTCTTGAAGCC 910
781 ACTGATCTTGCACCTTGAAGAGTTCTTGGAGAAATCCCATGAGTTTCAATGTTTGTG 840
911 CTTGACCTTGCACCTTGAAGACTTCTTGGAGAAATTCCTAATGATTTGTC 970
841 ATTCTCACTCCCAAGGATCTTGGCTCAAGACAAATGTTTGGGGTATCCCGACACCGGT 900
971 ATTCTTCTCCCAAGGATTTACTTGGCCCAAGATTAATGCTTGGAGATACCCCTGACCTGGT 1030
901 GGGCAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTTGAGAAATGAGATCTCCTCGT 960
1031 GGGCAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTTGAGAAATGAGATCTCCTCGT 1090
961 ATAAAGCAACAAGATCTCAACATCACCCCTGGAATCTCATTAATTAATGACCTTCTCT 1020
1091 ATTAAGCAACAAGATGGAATGATCTGATCTGATATCTCATTAATCACCCGCTCTCTCCC 1150
1021 GATGCTGTGGAACAACAATGCGGTCAAGACTTGAAGAAATGAGAAACAGAGCATCTCG 1080
|||||

1151 GATCAATTCGAACTACTTGTGGCCAAAGTCTTGAGAAAGTGTGCGAACCGAGCATCC 1210
1081 GATATTTCTGAGATACCTTTCAGAAACAGAAATGTTGTTCCAAAATGATCTCAGA 1140
1211 CACTTCTGAGATTCCTTTTGAATCTGAGAAAGGAATTTGTTCCGTGATGATCTCAGA 1270
1141 TTGAAAAGTCTGGCAATCTTGGAAAACCTACACAGAGATGTTGCTCATGAATCTCC 1200
1271 TTCC---AAGTCTGCAATCTTGGAAAACCTACAGAGATGTTGCTCATGACCTGGCC 1327
1201 AAAGATTCACGGCCAGCCAGATCTGATCATTCGGAACNACAGCGACCGGAATATGCTC 1260
1328 AAAGATTCACGGCCAGCCAGATCTGATCATTCGGAACNACAGCGACCGGAATATGCTC 1387
1261 GCTCTTGTGTCGACATTAATTTAGTGTGTCACAGTGCACCATCGCCATGCTTTGGAG 1320
1388 GCTCTTGTGTCGACATTAATTTAGTGTGTCACAGTGCACCATCGCCATGCTTTGGAG 1447
1321 AAGCAAAATATCCAGATTCAGATATCTATTGGAAGAGCTTGAAGCAAAATACATTC 1380
1448 AAGCAAAATATCCAGATTCAGATATCTATTGGAAGAGCTTGAAGCAAAATACATTC 1507
1381 TCTTGCCAAATTAAGCTGATCTTTTTCGAAATGAACATACAGATTTTCAATGACAGT 1440
1508 TCTTGCCAAATTAAGCTGATCTTTTTCGAAATGAACATACAGATTTTCAATGACAGT 1567
1441 ACTTTCAGGAATTTGAGAGAGAGAGAGACACTTGTGTCATATGAGAGCAGACTGCT 1500
1568 ACTTTCAGGAATTTGAGAGAGAGAGAGACACTTGTGTCATATGAGAGCAGACTGCT 1627
1501 TTCACTCTTCTGCTCTCTACCTGTTGTTGATCAGTATCGATGTTGATCCCAATTC 1560
1628 TTTCACCTTCTGCTCTCTACCTGTTGTTGATCAGTATCGATGTTGATCCCAATTC 1687
1561 AACATGTTTCCCGTGGTGCATATGAGATATACCTTCCCTTACACCGAAGAGAGCG 1620
1688 AACATGTTTCCCGTGGTGCATATGAGATATACCTTCCCTTACACCGAAGAGAGCG 1747
1621 AGTTGGAAGATTTCCATCTGAGATGAGAGACCTTCTTTCACCAAAAGTTGAGATGAA 1680
1748 AGTTGGAAGATTTCCATCTGAGATGAGAGACCTTCTTTCACCAAAAGTTGAGATGAA 1807
1681 GAACATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1740
1808 GAACATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1867
1741 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1800
1868 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1927
1801 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1860
1928 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1987
1861 GAGAGGCTGAAATGAGAAATTTTGAAGTGTGATGACAGATTAACATTTGAACGCGCA 1920
1988 GAGAGGCTGAAATGAGAAATTTTGAAGTGTGATGACAGATTAACATTTGAACGCGCA 2047
1921 TTTCAGATGATATCATCTCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1980
2048 TTTCAGATGATATCATCTCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2107
1981 TGCGACACGAAAGGTGCTTTGTACAGCTGATGATGATGATGATGATGATGATGATGAT 2040
2108 TGCGACACGAAAGGTGCTTTGTGACAGCTGATGATGATGATGATGATGATGATGATGAT 2167
2041 GTGAGGCAATGACTTGGCTTTGTCACATTTGCAACTGTAACGTTGAACGCGGAG 2100
2168 GTGAGGCAATGACTTGGCTTTGTCACATTTGCAACTGTAACGTTGAACGCGGAG 2227
2101 ATTAATGTTGCAATGAGAAATCTGTTTCAATGATGATGATGATGATGATGATGATGAT 2160
2228 ATCAATGTTGCAATGAGAAATCTGTTTCAATGATGATGATGATGATGATGATGATGAT 2287
|||||

QY 2161 GACATACGTCGATTTCTTTGAAAAGTGTAAAGAAAGTCCATCTCACTGGGATTAAGTTC 2220
DB 2288 GATCTCTCTGTTGATCTTTTGAAGAGTCAAGCTTGAACCCATCCGAGGAAACATC 2347
QY 2221 TCCCAAGAGGCTTGAACGAAATGAGAGAGATATACATGAAGATTATCTCGAGAGA 2280
DB 2248 TCAAGAGCTGGTCTCCAGCGATTAAGAGAAATGACATGCAAAATTTACTCAAGAG 2407
QY 2281 CTATTGACCTGACAGAGATATGATTTCTGAAAGCATTTTCCAACTTGAACGCCCT 2340
DB 2408 CTCTCACTCTCACTGGTGTCTATGAGCTTCTGGAAGCATGTGTCTAACTTGAACGCCCT 2467
QY 2241 GAGAGTGTGCTTCTCTTGAAGATTTTATGCTCTTAAGTACCGTAAAGCTGCTGAATCA 2400
DB 2468 GAGAGCGCGCTATCTCGAGATTTCTATGCTCTCAAGTACCGCAAAATTTGGCGAGCT 2527
QY 2401 GTTCATTTGACAGAGAGATTAATTGAACCTG 2431
DB 2528 CTGCCCTCTGCTGTGAGTAACTGAGGATG 2558

RESULT 5
US-10-425-114-6460
; Sequence 6460, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6460
; LENGTH: 2772
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: 700565776_FLI
US-10-425-114-6460

Query Match 59.8%; Score 1570.4; DB 13; Length 2772;
Best Local Similarity 78.2%; Pred. No. 0;
Matches 1912; Conservative 0; Mismatches 527; Indels 6; Gaps 2;

QY 1 ATGGCTGAGCGTGTCTCTCACTCGCGTCCAGAGTCTCGTGAAGCTTTGATGAGACCTT 60
DB 53 ATGGCAATATCTCTTTGACACACTCTCACTCTTCCGAGAGGTTTGAAGAACTCTC 112
QY 61 CTGTCTCAGAGAGAGATTTTGGCTTGTCTCAAGATCGAGGGCAAGAAAAGA 120
DB 113 ACTGTGACAGAGAGAGATTTTGGCTTGTCTCAAGGCTTGAAGCGCAAGGGA 172
QY 121 ATTCTGCAACACATCAAAATTAATCTAGATTGAAGTATCCCTGAGAGAAAG 180
DB 173 ATCTGCAACATCAACAGAGTGTGAGAGTTTGAAGAAATCCCTGAGAGAAAG 232
QY 181 AAGCTCGTAATGAGTATTTTGAAGTATGAAGCTAGTCAAGAAAGCATCGTTG 240
DB 233 AAATCTCCAAATGAGTCTTTGAGAAAGTTTGAATCACAAGAAAGCATAGTCTG 292
QY 241 CCTCAATGAGTGTGATTTGCTGTCTGCAAGCGCTGTGTTGGAGTATTAAGAGT 300
DB 293 CCACATTTTGAAGTCTGTGCTGCAAGCGCTGTGATGAGAAATCTCGTGTG 352
QY 301 AATGTTCAGCGCTGTGTTGAGAACTCACTGTTGCTGAGTATCTCACTTCAAGGA 360

DB 353 AATGTGACATGCTGTGTGTGATGAGCTTGTCTGCTGAGTATCTGCTTTCAAGAG 412
QY 361 GAGCTGTGATGAGAGTTCAAAATGAAAATTTGTTGAAATTTGATTTGAGACCTTC 420
DB 413 GAGCTGTGATGAGAGTTCAAAATGAGAACTTGTGTGATTTGAGCTTTGAAACCTTT 472
QY 421 AACTCATATTTCCCGCCCACTCTTCAAAAATGATGATGATGATGATGATGATGATGAT 480
DB 473 AATGATCTCTTCCCTGCGCACTCTGACAGATGATGATGATGATGATGATGATGATGAT 532
QY 481 AATGTGACCTTTGCGCAAAATTTGTTGATGACAGAGAGATGACACCTTGTCTGAA 540
DB 533 AATGCGCACTTTGCGCAAGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 592
QY 541 TTCTCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 593 TTCTCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
QY 601 TTGATGCTCTTCAACATGTTTGAAGAGAGAGAGATGATGATGATGATGATGATGATGAT 660
DB 653 CTGATCTCTCTCAACATGTTTGAAGAGAGAGAGATGATGATGATGATGATGATGATGAT 712
QY 661 GAGACACATGTCGCAATTCGAACACCGGTTCCAGAAATCGTTTGAAGAGAGTTG 720
DB 713 GAAACACCTTACTCGCAATTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 772
QY 721 GGTGACACCGCAAGAGCGGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 773 GGTGACACCGCAAGAGCGGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 832
QY 781 ACTGATCTCTGACACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 833 CTGACACCTTGAAG 892
QY 841 ATTCTCACTCCCAAG 900
DB 893 ATCTCTCTCCCAAG 952
QY 901 GGGAGAGTTGTTTCAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 953 GGAAGAGTTGTTTCAATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1012
QY 961 ATAAAGCAAG 1020
DB 1013 ATCAAGAAAG 1072
QY 1021 GATGCTGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1073 GATGCTGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132
QY 1081 GATGCTGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1133 GATGCTGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
QY 1141 TTGAAAG 1200
DB 1193 TTGAAAG 1249
QY 1201 AAG 1260
DB 1250 AAG 1309
QY 1261 GCTCTCTGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320
DB 1310 GCTCTCTGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1369
QY 1321 AAGCAAAATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1370 AAGCAAAATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
QY 1381 TCTTGCAATTTACAGTGTATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1430 TCTTGCAATTTACAGTGTATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489

1441 ACTTCCAGAAATTGACAGAGCAAGGACACTGTTGTCATAAGAGAGCCACACTGCT 1500
1490 ACCTTCCAGAGATTGCTGAGAGCAAGGACACTGTTGAGAGATAGAGATCACTGCCC 1549
1501 TTCACTCTCTGCTGCTTACCGGTTGTACATGCTATCGATGCTGATGCTTCCAAATTC 1560
1550 TTCAACCTTCCAGAGACTTACCGGTTGTATCGATGCTATGCTTGAATTCGCAAGGCTT 1609
1561 AACATGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTCAACCGAAGAGAGCGG 1620
1610 AACATGTTCTCTCCCGGTCGAGACATGAGTATATCTTCCATACATGCAAGTGAAGGT 1669
1621 AGTTGAAGCAATTTCCATCTGAGATGCAAGACCTTCTTTACACCAAGTTGAGATGAA 1680
1670 AGTTGAAGCAATTTCCATCTGAGATGCAAGACCTTCTTTACAGCTGAGTGAAGATGAG 1729
1681 GAACACTATGCTGCTGATGACCGGCAACGCAATCTGCTTCAACATGCAAGGCTT 1740
1730 GAACACTATGCTGATGAGAGACCGCAACCAATTAATCTTCAACATGCAAGGCTT 1789
1741 GATCGTCTCAAGAACTTAACCGGACTCGTCAAGTGTGCGGCAAGAACCCAAAGTGGCT 1800
1790 GACCTGTGAAGAAACATCAACGGGCTTGTGAGTGTATCGGGAAGAACGCAAGCTCCG 1849
1801 GAGTTGGCTTAACTCTGATGTTGATGTTGATGAGTGAAGAAAGAAATCTTAAAGATTGAA 1860
1850 GAGTTGGTGAAGCTGCTGCTGCTGCTGAGACAGAGAGAGAGAGTCAAGAACTTGGAA 1909
1861 GAGAAAGCTGAAGAAAGAAAGAAAGTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
1910 GAGAAAGCTGAAGAAAGAAAGTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1969
1921 TTCAAGATGATATCTCAATGAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
1970 TTCAAGATGATATCTCAATGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2029
1981 TGCAACAGAAAGGCTGCTTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040
2030 TGTAACAGAAAGGCTGCTTGTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2089
2041 GTGAGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100
2090 GTTGAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2149
2101 ATTAATGCTGAGAGAAATGCTGTTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2160
2150 ATCAATGCTGAGAGAAATGCTGTTTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2209
2161 GACATATCTGCTGATGTTTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220
2210 GAGATCTCTGTTGAGTCTTGTGAGAAAGAGCAAGGCTGATCTCACTGAGGAGCAAAATC 2269
2221 TCCCAAGAGAGCTGAAACGAATAGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2280
2270 TCCCAAGAGAGCTGAAACGAATAGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2329
2281 CTATTTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2330 CTCTTTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2389
2341 GAGAGTCTGCTTACCTTGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
2390 GAGAGCAAGCTTACCTGAGATGTTCTATGCTCTCAAGTACCGCAATTTGGCTGAGTCT 2449
2401 GTTCCATTTGGC---AGAGAGTGAATGACCTGTTAAATCAAT 2442
2450 GTGCCCTTGTCTATTTGAAGAGTGAATCATGTTTGAAGAGAAAT 2494

RESULT 6
US-10-424-599-11458
Sequence 11458, Application US/10424599

Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21 (53223) B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 285684
SEQ. ID NO 11458
LENGTH: 3360
TYPE: DNA
ORGANISM: Glycine max
FEATURES:
OTHER INFORMATION: Clone ID: PAT_MRT3847_110354C.1
US-10-424-599-11458
Query Match 59.8%; Score 1570.4; DB 13; Length 3360;
Best Local Similarity 78.2%; Pred. No. 0;
Matches 1912; Conservative 0; Mismatches 527; Indels 6; Gaps 2;
1 ATGCTGAGGCTGCTCACTCGGCTCCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
172 ATGCAATATCATCTTTGACACACTCTCACTTTCCGCAAGGTTTGAATGAATCTTC 231
61 CTGCTCAGAGAAAGAGATTTTGGCTTGTCTCTCAAGATGAGAGGCAAGAAAGAA 120
232 ACTGTCACAGAAAGAAATTTGGCTTGTCTCTCAAGGCTTGAAGGCAAGGCAAG 291
121 ATTCTGCAACCAATCAATTAATTTAGAGTTGAAGCTATCCCTGAAGAAAGAAAG 180
292 ATCTGCAATATCAACCAAGGTTGAGAGTGAAGAAATCCCTGAAGAGAGAAAG 351
181 AAGTCTGATATGATGATTTTGAAGTATGAGGCTGATGAGAAAGCATGCTG 240
352 AATCTCAAGATGATGATTTTGAAGTATGAGTCAACAGAGAGCATGATGCTG 411
241 CCTCATGAGGTTGACATTTGCTGCTGCTCAAGGCTGCTGTTGGAGTATCATGAGTG 300
412 CCACATTTGATGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
301 AATGTCACGCTTGTGTTGAGAACTGATGCTGATGATGCTGATGATGCTGATGATG 360
472 AATGTCACAGTGTGTTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
361 GAGTGTGATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 420
532 GAGTGTGATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 591
421 AACTCATATTTCCCGCCCAACTCTTCAAAATCCATGATGATGATGATGATGATGATG 480
592 AATGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
481 AATGTCACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
652 AATGTCACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
541 TTCTCAGAGTCCATTTGCTCAAGGCAAGATGATGATGATGATGATGATGATGATGATG 600
712 TTCTCAGAGTCCATTTGCTCAAGGCAAGATGATGATGATGATGATGATGATGATGATG 771
601 TTGAATGCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
772 CTGATTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 831
661 GAGACACATGCTGCTGATTTGAAACCGGTTCCAGAAATCGTTTGAAGAGTGG 720
832 GAAACACCTTACTGGAATTTGAAACCGGTTCCAGAAATCGTTTGAAGAGTGG 891
721 GTGACACCGGCAAGCGGCTGCTGAGATGATCACTCTTTGATCTTTTGAAGCA 780


```

Db      892  GGTACACCTGCGAGCGGTGCTCGAGATGATCCAGCTTCTCGGACCTTCTTGAGGCA 951
Qy      781  ACTGATCTTGGACACCTTGGAGAGTTCTTGGAGAAATCCCAAGGTGTCAATGTGTG 840
Db      952  CCTGACCTTGGACACCTTGGAGACATCTTGGAGAGTCCCTAATGTGTGT 1011
Qy      841  ATTCTCATCCCGACGAGTACTGCTGACAGACATGTTTGGGGTATCCGACACGGGT 900
Db      1012  ATCTCTTCTCCCAAGTGTACTTGGCCAGATATGTTCTTGGATACCTGACCTGGT 1071
Qy      901  GGGCAGGTTGTTTATCATCTTGGATCAAGTCCGAGCTTGGAGAAATGAGATCTCTCGT 960
Db      1072  GGACAGGTTGTTTATCATCTTGGATCAAGTGTGTGCTTGGAGAAATGAGATCTCAACGC 1131
Qy      961  ATAAAGCAACAAGATCAATCACTACCCCTGAAATCTCATTTATTAAGTCTTCTCT 1020
Db      1132  ATCAAGAAACAAGGCTTGAATATCACTCCCTGTATCTCATTTATCACTGCTCTCTCT 1191
Qy      1021  GATGCTGCGGAAACAATGCGGTCAAGATGAGATGAGAAATGAGAAAGATGAGACGACTCG 1080
Db      1192  GATGAGATGAGAACTACCTGAGCCAGCTCTAGAGGGTATATGATATCTGAATATGT 1251
Qy      1081  GATATCTTCCAGTACCTCTGAGAACAGAAAGGAAATGTTCCGAAAATGAGATCTCAGA 1140
Db      1252  GACATCTCAGAGTCTCTTCAAGAACAGAAAGGAAATGTTCCGAAAATGAGATCTCAGA 1311
Qy      1141  TTTGAAAAAGTCTGGCCATCTAGTGAATCTTACACAGAGATGTTCTCATGAATCTCC 1200
Db      1312  TTCC---AGTCTGGCCATCTAGAGACTTACACTGAGAGATGTCGCTTGAACCTGGCC 1368
Qy      1201  AAAGAGTGCAGCGGACGCGAGATCTGATCTCGGAAACGAGGAAATGAGATGCTC 1260
Db      1369  AAGAGTGCAGCGGACGCGAGATCTGATGTTGAGAACTAGAGAGGAAATGAGATGCTC 1428
Qy      1261  GCCTCTTGTCTGCAATAAATGAGTGTCAACAGTGCACCATGCGCCATGCTTGGAG 1320
Db      1429  GCCTCTTGTGAGCAATAAATGAGTGTCACTGAGTGTCACTGCTGCTGAGAA 1488
Qy      1321  AAGACAAAATATCCAGATTCAGATATCTATGGAAGAACTTGAAGACAAATACCATTC 1380
Db      1489  AAGACAAAATATCCAGATTCAGATATCTGAGAAATTTGAAGAAATATCATATTC 1548
Qy      1381  TCTTGCCTTCAATGAGCTGATCTTGTGCAATGAAACCATAGATTCATATCATCCAGT 1440
Db      1549  TCATGCAATTTACGCTGATCTTGTGCAATGAAACCATAGATTCATATCATCCAGC 1608
Qy      1441  ACTTCCAGAAATGTCAGAGAGCACTGTTGTCTAATACGAGGCACTGCT 1500
Db      1609  ACCTTCCAGAGATGTCGAGAGAGCACTGTTGACAGTATGAGAGTCACTGCTCC 1668
Qy      1501  TTCACTCTTCTGCTCTCTACCGTGTGTACAGTATGATGATGATCCCAATTC 1560
Db      1669  TTCACTCTTCTGAGCACTTACCGTGTGTACAGTATGATGATGATCCCAATTC 1728
Qy      1561  AACATTTGTTTCCCTGCTGATGATGAGATATCTTCCCTTCAACCGAAGAAAGGG 1620
Db      1729  AACATTTGCTCTCCGCTGAGAGATGAGTATATCTTCCATACACTGAAACTGAGGT 1788
Qy      1621  AGTTTGAAGATTTTCATCTGAGATCGAAGACCTTCTTACACCAAGTTGAGATGAA 1680
Db      1789  AGTTTGAAGATTTTCATCTGAGATCGAAGACCTTCTTACAGTCTGAGAGATGAG 1848
Qy      1849  GAAACATATGCTGATTTGAAGAGCGCAACCAATTAATCTTCAACCATGGAAGGCTT 1908
Db      1741  GATGCTGCAAGAACTTAAACCGGACTGCTGAGTGTGCTGCGCAAGAACCAAGTTGCT 1800
Qy      1909  GACGCTGGAAGAAATCACTGAGGCTTGTGAGTGTGAGGAGAAAGCAAGCCTCGC 1968
Db      1801  GAGTTGCTAACTGCTGATGTTGAGTGTGATAGGCAAGAAATTAAGATTTGGAA 1860

```

```

Db      1969  GAGTTGTGAACCTGTGTGTGTGCTGAGACAGAGAGAGAGTGAAGACTTGGAA 2028
Qy      1861  GAGAGGCTGAATTAAGAAATGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1920
Db      2029  GAGAGGCTGAATTAAGAAATGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2088
Qy      1921  TTCAGATGATATCTCAATTAAGAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
Db      2089  TTCAGATGATATCTCAATTAAGAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2148
Qy      1981  TGCACACGAAAGGCTTGTGACACCTGCTGATGATGATGATGATGATGATGATG 2040
Db      2149  TGTGACACGAAAGGCTTGTGACACCTGCTGATGATGATGATGATGATGATGATG 2208
Qy      2041  GTGAGGCAATGATCTGCGGTGTCGCAATGTCGCAATGTCGCAATGTCGCAATGTC 2100
Db      2209  GTTGAAGCAATGATCTGCGGTGTCGCAATGTCGCAATGTCGCAATGTCGCAATGTC 2268
Qy      2101  ATTATGCTGATGAGGATCTGCTTCAATGATCTTCAATGATCTTCAATGATCTTCA 2160
Db      2269  ATCATGCTGATGAGGATCTGCTTCAATGATCTTCAATGATCTTCAATGATCTTCA 2328
Qy      2161  GACATCTGCTGATCTTCTTGAAGAGTGAAGAGATGATGATGATGATGATGATGAT 2220
Db      2329  GAGATCTGCTGATCTTCTTGAAGAGTGAAGAGATGATGATGATGATGATGATGAT 2388
Qy      2221  TCCCAAGAGGCTTGAAGAGTGAAGAGATGATGATGATGATGATGATGATGATGAT 2280
Db      2389  TCCCAAGAGGCTTGAAGAGTGAAGAGATGATGATGATGATGATGATGATGATGAT 2448
Qy      2281  CTATTTGACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db      2449  CTCTTGAACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 2508
Qy      2341  GAGAGTGTGTACTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db      2509  GAGAGTGTGTACTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2568
Qy      2401  GTTCCATTTGCT---AGAGAGTAAATGAACCTGTTAAATGAAT 2442
Db      2569  GTTCCCTTGTCTAATGAAGATGAATTAATGATTTGAAGAGAACT 2613

RESULT 7
US-10-424-599-11459
; Sequence 11459, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 11459
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110355C.1
US-10-424-599-11459

Query Match          59.7%; Score 1568; DB 13; Length 3438;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1902; Conservative 0; Mismatches 536; Indels 3; Gaps 1;

Qy      1  ATGGCTGAGGCTGCTCACTGCGGTCCAGAGTCTCGGTGAGCGTTGATGAGACCTT 60
Db      689  ATGGCAATACCTTGAACACACTCTCACTCTTCCGAGAGGTTGATGAATCTCTC 748

```

Qy 61 CTGCTCAGAGAGAGATTGCGCTTCTCAGAGATCGAGGCAAGAAAAGGA 120
Db 749 ACTGGTCAAGAGAAATGAAATTTGGCCCTTTTGTCTCAAGCTTGAAGCCAAAGGCAAGGA 808
Qy 121 ATTCTGCAACCCANTCAATTAATCTAGATTGAAAGCTATCCCTGAAGAGACAGAAG 180
Db 809 ATCTGCAACCAACAGAGTGGTGCAGAGTTTGAAGAAATCCCTGAGAGACAGCAAGAG 868
Qy 181 AAGCTCGCTAAATGATGCAATTTTGAAGTATGAAGCTAGTCAAGAACCATGCTTTG 240
Db 869 AAATCTCAAGAGTGGTGTCTTTGAGAAATTTTGAAGATCTACAGAGAACCATAGTGTG 928
Qy 241 CCTCCATGGGTTGCACTTGTCTGTTCGCAAGGCTGTGTGTTGGAGTACATTAAGTGTG 300
Db 929 CCACCATTTGGGCTGTGGCTGTTCGACCAAGGCTGTGTGTTGGAGTATCTCGGGGTG 988
Qy 301 AATGTCCAGGCGCTGTGTTGAGAGACTCACTGTGTGAGTATCTCCACTTCAGAGAA 360
Db 989 AATGTGCACATGCTGTGTGTGATGAGCTGTCTCTGCTGAGTATCTGCTTTCAAGAG 1048
Qy 361 GAGCTGTGATGAGAGTTCAAAATGAAACTTGTGTTGAAATGAAATTTGAGACCTTC 420
Db 1049 GAGCTGTGAGAGAAAGTTCTAAATGCACTTTGTGCTTGAAGTGAATTTGAACCTTT 1108
Qy 421 AACTCATATTCCCGCCCACTCTTCAAAATCATGTGTAAATGTGTGAGTCTTA 480
Db 1109 AATGATCTCTTCCCTCGCCCACTCTGAACAAATCATGTGAAATGTGCTGAGTCTTC 1168
Qy 481 AATGTCACTTTGGGCAAAATGTTTGCATGACAAAGAGACATGACCTTTGCTGGA 540
Db 1169 AATGCACTTTGGGCAAGCTCTTCCATGACAAAGAGACATGACCTGCTGTGGA 1228
Qy 541 TTCTCAGAGTCAATGTCAAGAGGCAAGAACATGATGTGAATGACAGAAATTCAGAAC 600
Db 1229 TTCTCAGAGCTTCAAGTTATTAAGGAAAGAACATGATGTGAATGACAGAAATTCAGAAC 1288
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGATCTTGTGATCCCTTACTCT 660
Db 1289 CTGGATCTCTCCAGCATGTTTGAAGAAAGCAGAGATCTGATCTTCACTGATGCTCT 1348
Qy 661 GAGACACATGCGGCAATTTGAAACACCGGTTCCAGGAAATCCGTTTGAAGAGGTTGG 720
Db 1349 GAAACACCTTCTCAAGTATGAAAAAATTCGAGAAATTTGTTGGAAGAGGAGGTTGG 1408
Qy 721 GGTGACACCGGAGAACGCGTCTCGAGATGATCCAACTCTTTTGGATCTTCTGAGGCA 780
Db 1409 GGTGACATCGCGAGCGTGTCTCGAGATGATCCAACTCTTCTTGAACCTTCTGAGGCA 1468
Qy 781 ACTGATCTTGAACCTTGAAGAGTCTTGTGAGAGATCCCAATGATGTTCAATGTTGTG 840
Db 1469 CCGGACCTTCAACCTCGAGACATTCCTTGAAGAGTCTTCAATGATGTTGTG 1528
Qy 841 ATTCTCACTCCCAACGGAATCTTGTGCAAGAAATGTTTGGGGTATCCGGAACCGGT 900
Db 1529 ATCTCTTCTCCCAAGTTACTTGTGCCAAGAAATGTTCTTGGGGTATCCGGAACCGGT 1588
Qy 901 GGGCAGGTTGTTTATCATCTTGGATCAAGTCCGAGCTTGTGAGAAATGAGANTCTCTCGT 960
Db 1589 GGAAGAGTTGTTTATCATCTTGGATCAAGTTCGTGCTTGTGAGAAATGAGATCTCAACGC 1648
Qy 961 ATAAAGCAAGAGCTCAACATCAACCCCTCGAATCTCATTAATTAAGACTTCTTCT 1020
Db 1649 ATCAAGAAACAAGGCTGATATCAACCCCTCGAATCTCATTAATTAAGACTTCTCTCT 1708
Qy 1021 GATGTGTGAGAAACAATGTGGGTCAAGACTTGAAGAAATGAGAAAGAGACACATCG 1080
Db 1709 GATGTGTGAGAAACAATGTGGGTCAAGACTTGAAGAGGATATGATATCAAGATATGT 1768
Qy 1081 GATATCTTGTGAGTACCTTCAAGACAGAAAGGAAATGTTTCAAAATGATCTCAGA 1140
Db 1769 GATATCTTGTGAGTACCTTCAAGACAGAAAGGAAATGTTTCCCAATGATCTCAGA 1828

Qy 1141 TTTGAAAAGCTGTGCGCATACTTGGAAACCTTACACAGAGGATGTGCTCAAGAAATCTCC 1200
Db 1829 TTGG--AACTGTGGCCATATCTTGAAGACTTGAAGCTTGAAGATGTTGCTTGAACCTGCC 1885
Qy 1201 AAAGAGTTGACAGGCAACGCGAGATCTGATCATCGSAAACAAACAGCGCAATATCTGC 1260
Db 1886 AAGAGTTGCAAGCCAAAGCCAGATCTGATGTTGAAACCTACAGTGTATGAAACATTTGT 1945
Qy 1261 GCTCTCTTGTCTGCAATAAATTAAGTGTGACACAGTGCACCATGCGCATGCTTGGAG 1320
Db 1946 GCTCTTGTGTTAGACATAAATTAGAGTAACTCAGTGTACCATGCTCATGCTCTGAA 2005
Qy 1321 AAGCAAAATATTCAGATTTAGATATTAATGGAAGAGCTTGAAGCAATATCAATTC 1380
Db 2006 AAGCAAAATATTCAGATTTAGATATTAATGGAAGAGCTTGAAGCAATATCAATTC 2065
Qy 1381 TCTTGCAATTTACAGTATCTTTTGAATGAAACCATACAGATTTATCATCAACAGT 1440
Db 2066 TCAATGCAATTTACTGTGATCTTTTGAATGAAACCATACAGATTTATCATCAACAGC 2125
Qy 1441 ACTTCCAGGAAATGCAAGAGCAAGACATGTTGTCAATACAGAGCCACATGCT 1500
Db 2126 ACTTCCAGAGATGCTGTGAAGCAAGACATGTTGACAGTATGAGATGACACTGCC 2185
Qy 1501 TTCACTTTCTGTGTCTTACCGTGTGTGATATGATATGATGTGTTGATCCCAATTC 1560
Db 2186 TTCACTTTCTGTGTGTGATATGATATGATATGATATGATATGATATGATATGATATG 2245
Qy 1561 AACATGTTTCCCTGTGTGTGATATGATATGATATGATATGATATGATATGATATGAT 1620
Db 2246 AACATGTTTCCCTGTGTGTGATATGATATGATATGATATGATATGATATGATATGAT 2305
Qy 1621 AAGTTGAAGATTTTCACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 2306 AAGTTGAAGATTTTCACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2365
Qy 1681 GAACACTATGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 2366 GAACACTATGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2425
Qy 1741 GATGTGTCAAGAACTTCAACCGGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1800
Db 2426 GATGTGTCAAGAACTTCAACCGGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2485
Qy 1801 GAGTTGTCAACCTCGTGTGTGTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGAT 1860
Db 2486 GAGTTGTCAACCTCGTGTGTGTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGAT 2545
Qy 1861 GAGAGGCTGAATGAGAAATGTTTGAAGTGTGATGATGATGATGATGATGATGATGAT 1920
Db 2546 GAGAGGCTGAATGAGAAATGTTTGAAGTGTGATGATGATGATGATGATGATGATGAT 2605
Qy 1921 TTCAATGATATCATCTCAATGAAAGAGATCCGAATTTGAACCTTACCGTATCAT 1980
Db 2606 TTCAATGATATCTCTCTCAAGTGAACCGTGTGAAGAACGAGAGCTCTACCGTATCAT 2665
Qy 1981 TGCACACGAAAGGTGCTTTGTGACAGCTGTGATGTGATGAGCTTTGATTTGACGTT 2040
Db 2666 TGCACACGAAAGGTGCTTTGTGACAGCTGTGATGTGATGAGCTTTGATTTGACGTT 2100
Qy 2041 GTGAGGCAATGATCTGCGGTTTGTGCAACATTTGCAACCTGTAAACGTTGAACGCGAG 2100
Db 2726 GTGAGGCAATGATCTGCGGTTTGTGCAACATTTGCAACCTGTAAACGTTGAACGCGAG 2785
Qy 2101 ATTAATGTCCATGGAATCTGTTTCAACATTTGATCTTACCATGAGTGTGATCAAGTGT 2160
Db 2786 ATTAATGTCCATGGAATCTGTTTCAACATTTGATCTTACCATGAGTGTGATCAAGTGT 2845
Qy 2161 GACATATCTGTGATTTCTTTGAAAGTGTGAAGAGATCATCTCACTGAGATTAAGATC 2220
Db 2846 GACATATCTGTGATTTCTTTGAAAGTGTGAAGAGATCATCTCACTGAGATTAAGATC 2280
Qy 2221 TCCCAAGAGGCTTGAAGCAATGAGAGAGATATGATGAGAGATTTACTCGAAGAGA 2280

Db 2906 TCCAGGCTGACCTCAGAGCTATTCATGAGAGTACATGAGCAATTAATCTGACAGG 2965
Qy 2281 CTATTGACCCGAGAGAGAGTATGAGATTCGAGAGCATGTTTCCAACTTGAACGGCT 2340
Db 2966 CTCTTGACATCAGCTGAGTGTATGAGCTTCTGAGAGCATGAGCAATTAATGAACGGCT 3025
Qy 2341 GAGAGTCGCTGATCTTGAAGTGTATGAGCTTCTGAGAGCATGAGCTGAGTATCA 2400
Db 3026 GAGAGCAACCTTACCTTGAAGTGTATGAGCTTCTGAGAGCATGAGCTGAGTATCA 3085
Qy 2401 GTTCATTCGAGAGAGTAAATGAGCTTGAATTAACA 2441
Db 3086 GTGCCCCCTTGCTATGAGAGTAAATCATGTTTGAAGAGA 3126

RESULT 8
US-10-425-114-9594Sequence 9594, Application US/10425114
Publication No. US2004003488A1

GENERAL INFORMATION:

APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 9594
LENGTH: 2772
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700852649_FLI

US-10-425-114-9594

Query Match 59.7% Score 1566.4; DB 13; Length 2772;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1901; Conservative 0; Mismatches 537; Indels 3; Gaps 1;

Qy 1 ATGCTGAGCGTCTCTCACTGCGCTCAGAGCTCTCGTGAAGCTTGGATGAGACCTT 60
Db 53 ATGCAATCAACCTTTGACACACTCTCACTCTTCCGAGAGGTTGATGAATCTCTC 112
Qy 61 CTGCTCAGAGAAAGAGATTTGGCTTGTCTCAAGATCGAGGCAAGAAAGAA 120
Db 113 ACTGTCACAGAAATGAATTTGGCTTTGTCTCAAGCTTGAAGCAAGGCAAGGAA 172
Qy 121 ATTCGCAACACATCAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGACAGAA 180
Db 173 ATTCGCAACACACATCAAGTGTGTCAGAGTTGAAGAAATCCCTGAAGAGAGAGAA 232
Qy 181 AAGCTCGTAATGTCATTTTGAAGTATGAAGCTAGTCAGAGAGCGATCTGTTG 240
Db 233 AAATCCCAAGGTGTCTTGAAGATTTGAATCTACACAGAGAGCAATAGTCTG 292
Qy 241 CCGTCAAGGTTGACCTTGTCTGTCAGAGAGCTGAGTTTGGAGATCAATTAAGATG 300
Db 293 CCAACATTTGAGCTCTGCTGCTGTCAGACAGAGCTGAGTTTGGAGATCTGCGAGT 352
Qy 301 AATGTCACGCTTGTGTGAGAGACATCTGTCGATATCTCACTTCAAGAA 360
Db 353 AATGTCACATGCTTGTGTGAGAGCTCTGTCGATATCTGCTTCAAGAG 412
Qy 361 GAGCTTGTGAGAGAGTCAATGAAGCTTGTGTGAGATGAGATTTGAAGCTTTC 420
Db 413 GAGCTTGTGAGAGAGTCTAATGAGCACTTGTGTGAGATGAGCTTGAAGCCGTT 472

Qy 421 AACTCATATTCCTCCGCGCAACTTTTCAAAATCCATGGTAATGTGTGAGCTTCTTA 480
Db 473 AATGATTCCTTCCCTCGCGCAACTTGAACAGATTCATGGAATGTGTGAGCTTCTTA 532
Qy 481 AATGTCACCTTTGGCAAAATTTGTCATGACAGAGAGATGACACCTTTGCTGAA 540
Db 533 AACGCAACCTTTGGCGCAAGCTCTTCAATGACAGAGAGATGACAGCACTCTTGA 592
Qy 541 TTCTCAAGATGATGTCACAGAGAGCAAGATGATGAGATGACAGATTCAGAA 600
Db 593 TTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 652
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTGATCCCTACTCT 660
Db 653 CTGATTCCTTCAAGATGTTTGAAGAAAGAGAGATGATGATGATGATGATGATGAT 712
Qy 661 GAGACACCATGTCGAGATTTGAACACCGGTTCCAGAAATCGTTTGAAGAGCTTGG 720
Db 713 GAAACACCTTCACTGAGATTTGAAGAAACAAATTCGGGAAATGCTTGAAGAGGAGT 772
Qy 721 GGTGACACCGAGAGAGGCTGAGATGATGATGATGATGATGATGATGATGATGATG 780
Db 773 GTGACATGCGCGAGCTGCTGAGATGATGATGATGATGATGATGATGATGATGATG 832
Qy 781 ACTGATCTTGCACCTTGAAGAGTCTTGAAGAAATCCCATGATGATGATGATGATG 840
Db 833 CCGACCTTGCACCTTGAAGAGATCTTGAAGAGATGATGATGATGATGATGATGATG 892
Qy 841 ATTCCTACCTCCAGAGATCTTGTCTGACAGAAATGTTTGGGATTCGCAACCGGT 900
Db 893 ATCTCTTCTCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952
Qy 901 GGGGAGGTTTTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 953 GGAAGGTTTTCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
Qy 961 ATAAAGCAACAGAGATCAACATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1013 ATCAAGAAACAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1072
Qy 1021 GATGCTGCGAACAACATGCGTCAACAGCTTGAAGAAATGATGATGATGATGATG 1080
Db 1073 GATCAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132
Qy 1081 GATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1133 GACATTCCTCGAGTCTTTCAGAACAGAAAGGAAATGTTGCAATGATGATGATG 1192
Qy 1141 TTGAAAAAGCTGCGCATATCTTGAAGAACTTCAAGAGATGATGATGATGATGAT 1200
Db 1193 TTGG---AAGTGTGCGCATATCTTGAAGAACTTCAAGAGATGATGATGATGATG 1249
Qy 1201 AAAGATGTCAGGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1250 AAGAAATGTCAGGCAAGCATGATGATGATGATGATGATGATGATGATGATGATG 1309
Qy 1261 GCTCTCTGTCGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1310 GCTCTGTTGTCGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1369
Qy 1321 AAGCAAAATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1370 AAGCAAAATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429
Qy 1381 TCTTGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1430 TCAATGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
Qy 1441 ACTTTCAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1490 ACTTTCAG 1549
Qy 1501 TTCACTCTCTGCTCTCACTCGTGTGATGATGATGATGATGATGATGATGATGATG 1560

```

Db      1550  TTGACCCCTTCAGAGACTCTACCGTGTGTTCACGGTATGTATCCCTTGTGATCCAAAGTTTC 1609
Qy      1561  AACATGTTTCCCTGCTGATGATGAGATATATCTTCCCTTACACCGAAGAGACGG 1620
Db      1610  AACATGCTCTCTCCCGGTCGCGACATGAGCATATATCTTCCCATACATCGAAGACTGAGCT 1669
Qy      1621  AGTTGAAGCATTTTCATCTCGAGATCGAAGACCTTCTTAAACCAAGTTGAGATGAA 1680
Db      1670  AGTTTAAACAGAGTTTCAACCCCGACATTTGAAAGCTTCTTACAGCTCAGTGGAAATGAA 1729
Qy      1681  GAAACCTTATGTGCTGATGACCGGCAACAGCCAAATTCGTTCACAAATCCCAAGGTT 1740
Db      1730  GAAACATATGTGATTTGAAAGACCGCAACAGCCGATCTTCAACATGCAAGACTT 1789
Qy      1741  GATCGTGTCAAGAACTTAAACGSACTCGTCAGTGTGCGGCAAGAACCCAAAGTTGCT 1800
Db      1790  GACCGTGTGAAGAAACATCAACGGGACTTGTGAGTGTGTAATGGCAAGAAATCCCGCTCCG 1849
Qy      1801  GAGTTGCTTAACTCTGTAGTTGTAGTGTGATAGCGCAAGAAATCTTAAAGATTTGAA 1860
Db      1850  GAGTTGTAAACCTGTGTGTGTGTGCGCGAAGACAGAGAAAGAGTCCAAAGACTTGAA 1909
Qy      1861  GAGAAAGCTGAAATGAAATGAAATGTTTGAAGTGTGATCCGAACTTGAACGGCCAA 1920
Db      1910  GAGAAAGCGCGAGATGAAAGATGTATGCGCTCATCGAGACCTTCAAGTTGAACGGCCAA 1969
Qy      1921  TTCAGATGATATCATCTCAATAGACAGAAATCCGAAATGTTGAACCTTACCGATACAT 1980
Db      1970  TTCAGATGATATCTCTCTCAATAGAACCGTGTGAGAAACGAGAGCTTACCGTGTATC 2029
Qy      1981  TGCAGACGAAAGGTGCTTTGTACAGCGTGTGATGATGAAAGCTTGTGATGACAGTT 2040
Db      2030  TGTGACAAAGGGGTGCTTTGTGACAGCGTGTGATGAGCTTGTGAGCTTGTGATGAGTGTG 2089
Qy      2041  GTGAGGCAATGATCTTGCGGTTTGGCAATTTCCGAACTGTAACGGTGAACAGCCGAG 2100
Db      2090  GTTAAAGGCAATGATCTTGCGGTTTGCACAACTTTGCAATGATGATGATGATGATGATG 2149
Qy      2101  ATTAATGTCATGAGAAATCTGTTTCAACATTTGATCTTACCAATGATGATGATGATG 2160
Db      2150  ATCATTTGATGATGAAATCTGTGTGATGATGATGATGATGATGATGATGATGATGATG 2209
Qy      2161  GACATATCTGTGATTTCTTTGAAAGTGTAGAAAGATTCATCTCACTGGATATGATC 2220
Db      2210  GAGATCTCTGTGATTTCTTTGAAAGTGTAGAAAGATTCATCTCACTGGATATGATC 2269
Qy      2221  TCCCAAGAGGCTTGAACGAATAGAGAGATGATGATGATGATGATGATGATGATGATG 2280
Db      2270  TCCCAAGAGGCTTGAACGAATAGAGAGATGATGATGATGATGATGATGATGATGATG 2329
Qy      2281  CTATTTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db      2330  CTCTTTGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2389
Qy      2341  GAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Db      2390  GAGAGCAACGTTACTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2449
Qy      2401  GTTCCATTTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2441
Db      2450  GTGCCCCCTTGTATTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 2490

```

RESULT 9
US-10-393-840-139

; Sequence 139, Application US/10393840
; Publication No. US20030229922A1

; GENERAL INFORMATION:

; APPLICANT: Bloksberg, Leonard N.

; TITLE OF INVENTION: Materials and Methods for the

; FILE REFERENCE: 11000.1012c3

```

; CURRENT APPLICATION NUMBER: US/10/393,840
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 139
; LENGTH: 2906
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-393-840-139

Query Match      57.1%; Score 1499.4; DB 16; Length 2906;
Best Local Similarity 76.5%; Pred. No. 0;
Matches 1852; Conservative 0; Mismatches 567; Indels 3; Gaps 1;

Qy      1  ATGCTGAGGCTGCTCATCTCGGTCACAGTCTCGTGAGCGTTGATGAGACCTT 60
Db      233  ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 292
Qy      61  CTGCTGACAGAAAGAGATTTTGGCTTCTCTCAAGATGAGGCAAGAAAGAA 120
Db      293  TGTGCTACCGCAACGATATTTGTGCTTCTTCAAGGTTGAAAGCGCAAGG 352
Qy      121  ATTCTGCAACCATCAATTTATTTGAGTTTGAAGTATCTCTGAAAGAAACGA 180
Db      353  ATCTTGAAGGCAACCGATTTTGTGAGTTGAGGCAATCTCTGAGAGAGCA 412
Qy      181  AAGCTGCTAATGATGATTTTGAAGTATGAGGCTGTGAGAAAGGATGATG 240
Db      413  AAGCTTTGATGAGGCTTTGATGAGGCTTCTCAATCTCAAGAAAGGATG 472
Qy      241  CTTCCATGAGGTTGATCTGCTGTCTTCAAGGCTGTGTGAGAGTACATTA 300
Db      473  CTTCCATGAGGTTGATCTGCTGTCTTCAAGGCTGTGTGAGAGTACATTA 532
Qy      301  AATGTACGCTTGTGTGAGAACTGATGATGATGATGATGATGATGATGATG 360
Db      533  AATGTACGCTTGTGTGAGAACTGATGATGATGATGATGATGATGATGATG 592
Qy      361  GAGCTTGTGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 420
Db      593  GAGCTTGTGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 652
Qy      421  AACTCATATTCCTCCGCGCACTCTTCAAAATCCATGATGATGATGATG 480
Db      653  AACTCATATTCCTCCGCGCGCACTCTTCAAAATCCATGATGATGATGATG 712
Qy      481  AATGTACCTTTTGGGAAATTTGATGATGATGATGATGATGATGATGATG 540
Db      713  AATGTACCTTTTGGGAAATTTGATGATGATGATGATGATGATGATGATG 772
Qy      541  TTCTGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 600
Db      773  TTCTGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 832
Qy      601  TTGAATGCTTTCAACATGTTTGAAGAAAGAGAGATGATGATGATGATG 660
Db      833  GTGTTCTCTTCAACATGTTTGAAGAAAGAGAGATGATGATGATGATG 892
Qy      661  GAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db      893  GAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
Qy      721  GGTGACCGGAGAGGCTGATGATGATGATGATGATGATGATGATGATG 780
Db      953  GGTGACCGGAGAGGCTGATGATGATGATGATGATGATGATGATGATG 1012

```

Oy 781 ACTGATCTTGACCCCTTGAGAAAGTTCTTGAGAGAAATCCCAGTGTGTCATATTTG 840
 Db 1013 CCGAGCCGTGCACTCTGAGAAAGTTCTTGATAGAGTTCCCAATGCTTTCAACGTCGTG 1072
 Oy 841 ATTCTCATCCCAAGGATTAATCTGCTCAAGCAATGTTTTGGGTATCCGCAACCGGT 900
 Db 1073 ATCATGTCCTCCCAAGGATTAATCTTGCTCAAGCAAGTCCTTGATATCCGATATCCGGT 1132
 Oy 901 GGGCAGGTGTTTACATCTTGATCAAGTCCGAGCTTTGAGAAATGATGTCCTCCGT 960
 Db 1133 GGGCAGGTGTTTACATCTTGATCAAGTCCGAGCTTTGAGAAATGATGTCCTCCGT 1192
 Oy 961 ATTAAGCAAGAGCTCAATCACTCCCTGATATCTCAATTAATCTAGACTTCTTCT 1020
 Db 1193 ATTAAGCAAGAGCTCAATCACTCCCTGATATCTCAATTAATCTAGACTTCTTCTCA 1252
 Oy 1021 GATGCTGTGCAAGCAATGCGGTCAAGCACTTGAGAAAGTATACGAAACAGAGCACTCG 1080
 Db 1253 GACGCGTTGGAAACCACTGTGCGCAAGGCTTGAGAAAGTTTTGGGACCGAGTACTTC 1312
 Oy 1081 GATATTTCTTGAGTACCTTCAAGCAGAAAGGGAATTTGCGAAATGATCTCAAGA 1140
 Db 1313 CACATTTCTTGCGGTCCCTTCAAGAAATGAGAGGAGTGTCCGCAAGTGGATTTCCCGG 1372
 Oy 1141 TTTGAAAAGTCTGGCCATCTTGGAACCTACAGAGAGATGTTGCTCATGAATCTCC 1200
 Db 1373 TTCC--AGGTGTGGCCCTATTTTGAAGATACATGAGAGATGTCGACGCAACTTCTCT 1429
 Oy 1201 AAAGAGTTGACGCGACCGCAGATCTGATCATCGGAAACNACAGCGCACTATCTGTC 1260
 Db 1430 GAGAGTTGACGCGGAGGAGCTGATCTGATCATCGGAAACNACAGGATGAGAAACATTTGT 1489
 Oy 1261 GCTCTCTTGTGCGCATTAATTAAGTGTCAACAGTGCACCATGCCCCATGCTTTGAG 1320
 Db 1490 GCTTCTCTGTAGCATTAATTAAGTGTCAACAGTGTACAAATGCCCCATGCCCCGAG 1549
 Oy 1321 AAGACAAATATCCGATTTGATATCTATGAGAAAGCTTGAGAAAGAAATACATTTT 1380
 Db 1550 AAGACAAAGTACCCGAGTCAAGACATTAATGAGAAAGAAATTTGAGAAAGTACCATTC 1609
 Oy 1381 TCTTGCCAAATTTACAGTATCTTTTTCATATGAACATATGATTTTCATCATCCAGT 1440
 Db 1610 TCTTGCCAGTTTCACTGCTGATCTATCTCCCATGAACCAACGATCTTATATCAACG 1669
 Oy 1441 ACTTTCCAGAAATTTGAGAAAGCAAGCACTGTGTCAATACGAGACCACTGCT 1500
 Db 1670 ACTTTCCAGAAATTTGAGAAAGCAAGCACTGTGTCAATACGAGACCACTGCAAC 1729
 Oy 1501 TTTCACTCTCTGCTGCTCTACCGGTGTGTACATGATCTGATGTGTTGATCCCAATTC 1560
 Db 1730 TTTCACTCTCTGCTGCTCTACCGGTGTGTCAACGAGATGAGCTTTGACCCGAAAGTTC 1789
 Oy 1561 AACATTTGTTCCCGTGTGCTGATATGAGATATATCTTCCCTTACACGGAAGAGGCG 1620
 Db 1790 AACATTTGTTCAACGAGTGTGATGATGATCTTATCTTGTCTTACCGGAAGAGGCG 1849
 Oy 1621 AGGTGGAAGATTTCCATCTGAGATGAGAAAGCTTTCTTTACACCAAGTGTGAGATGA 1680
 Db 1850 CGGTGGAATCTTCCATCTGAGATGAGAAAGCTTTCTTTACAGGATGTGAGAAAG 1909
 Oy 1681 GAACTTATGTGTGCTATGATACGCAAGCAAGCAATTTGTTTCAATGCGAAGGCTT 1740
 Db 1910 GAACTTGTGTGTGTTGAATATGAGAAAGCTTATTTATTTTCAACATGCGCAAGGCTG 1969
 Oy 1741 GATCGTGTCAAGAACTTAACGCACTGTCGAGTGTGTCGCGGAGAAACCAAGTGTGCT 1800
 Db 1970 GACGTGTCAAGAACTTTGAGAGGCTTGTGAGTGTGTATGCGAAGAACTTCAAGTGTGAG 2029
 Oy 1801 GAGTGTGCTAACCTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 Db 2030 GAACTGCGCAACTGTGTGTGTGAGAGTGAAGAGAGATTTCAAGAGACTTGGAA 2089

Oy 1861 GAGAGGCTGAAATGAGAAATGTTGAGTATGACAGTACAACTTGAGCGCCAA 1920
 Db 2090 GAGAGCTGTGAGTGAAGAAATGATACGACTTACGAAAGTACAGAGTATGCGCAG 2149
 Oy 1921 TTCAATGATATCATCTCAATGACAGAAATCCGAAATGTTGAATTTTACCATATT 1980
 Db 2150 TTCAATGATATTTCTTCCAGATGAAACCGGTGAGAGTGTCTTACCGCTACATC 2209
 Oy 1981 TGGACAGAAAGTGTCTTGTACAGCTGATGATGATGAGAGCTTTGATGACATT 2040
 Db 2210 TGTACAGAAAGAGTCTTGTGTAACCGGCTATCTATGAGAGCTTTGAGAGCTG 2269
 Oy 2041 GTGAGGCAATGATCTGCGGTTTCCCAATCTGCAACTGTAAACGATGACAGCGAG 2100
 Db 2270 GTTGAAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2329
 Oy 2101 ATATGTCATGGAATCTGTTTGAACATGATCTTACATGATGATGATGATGATGAT 2160
 Db 2330 ATCATTTGCAATGCAATCGGCTACCACTATCTTACATGATGATGATGATGATGAT 2389
 Oy 2161 GACATCTCTGCAATTTCTTGAAGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2390 GAGCTTCTGTAGCTTCTTCAACAGTGAAGATGATGATGATGATGATGATGATGAT 2449
 Oy 2221 TCCCAAGAGGCTTGAAGCAATGAGAGATGATGATGATGATGATGATGATGATGAT 2280
 Db 2450 TCAAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2509
 Oy 2281 CTATTTGACCTGACAGAGTGTATGATTTGGAAGATGTTTCAACTTGAACGCGCT 2340
 Db 2510 CTGTTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2569
 Oy 2341 GAGAGTGTGCTTACCTTGAATGTTTATGCTTTATGATGATGATGATGATGATGAT 2400
 Db 2570 GAGAGTGTGCTTACCTTGAATGTTTATGCTTTATGATGATGATGATGATGATGAT 2629
 Oy 2401 GTTCCATTTGAGAGAGATTA 2422
 Db 2630 GTTCTCGGCTGTGAGATTA 2651

RESULT 10
 US-10-393-840-15
 ; Sequence 15, Application US/10393840
 ; Publication NO. US20030229922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bloksberg, Leonard N.
 ; TITLE OF INVENTION: Materials and Methods for the
 ; FILE REFERENCE: 11000.1012c3
 ; CURRENT APPLICATION NUMBER: US/10/393,840
 ; PRIOR FILING DATE: 2003-03-20
 ; PRIOR APPLICATION NUMBER: US 09/636,800
 ; PRIOR FILING DATE: 2000-08-10
 ; PRIOR APPLICATION NUMBER: US 09/170,862
 ; PRIOR FILING DATE: 1998-10-13
 ; PRIOR APPLICATION NUMBER: US 60/148,426
 ; PRIOR FILING DATE: 1999-08-11
 ; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 956
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15
 ; LENGTH: 2913
 ; TYPE: DNA
 ; ORGANISM: Eucalyptus grandis
 US-10-393-840-15

Query Match 57.1%; Score 1497.8; DB 16; Length 2913;
 Best Local Similarity 76.4%; Pred. No. 0;
 Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
 Oy 1 ATGCTGAGCGTGTCTCATCTCGGTCACAGTCTCGGTGAGCGTTTGATGAGACCTT 60


```
Db 242 ATGGGTGATTCAGTGTGACTCGAAGCCAGCCTTCCGAGCGTTTGAGAGAGCCCTC 301
Qy 61 CTTCCTCACAGAAACGAGATTTTGGCTTCGTCATGAGATCGAGGCAAGAAAGAA 120
Db 302 TCTGTCTACCCGACAGATATTTGTGCTTCCTTTCAGAGGTTGAAGCCAGGCAAGGC 361
Qy 121 ATTCTGCAACCACTCAAAATTAATCTAGAGTTGAAGCTATCCCTGAGAGAACAGAAAG 180
Db 362 ATCTTGACAGCCCAACAGATTTTGTCTGAGTTGAGGCCATCTCGAAGAGAGCAGAC 421
Qy 181 AAGCTCGGTATNGGTCATTTTGAAGTAATGAAGCTAGTCAGAGAGAGAGAGAGAG 240
Db 422 AAGCTCTTGAGAGGAGCTTTGGTGAAGTCTCAAAATCCTCAGAGAGAGAGAGAGAG 481
Qy 241 CCTCCATGGGTTGCACTTGTGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 482 CTTCAATGGGTTGCTCTGTCGTCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
Qy 301 AATGTTCAAGCCTTTGTTGTTGAGAGACTACTGTTGATGATCTCCACTTCAGAGAA 360
Db 542 AAGCTCAAGCGCTTGTCTTGAGCAATGAGAGTTGCTGAGATCTGCACTTCAAGAA 601
Qy 361 GAGCTTGTGATGAGATTCAGAAATGGAATCTTGTGGAATGGAATTTGAGAGCCCTC 420
Db 602 GAGCTTGTGATGAGAGCTGATGATGTAATTTGTGCTGAGCTTGAAGCTTCAGAG 661
Qy 421 AACTCATATTCCTCCGCGCAACTCTTTCAGAAATCATTTGGTAATGATGAGAGTCT 480
Db 662 ACTGCTCTTTTCCGCGCCGCACTCTTTCAGAGTATATGAGCAATGGGCTGAGATTC 721
Qy 481 AATGTCACCTTTCCGCAAAATTTGTCATGACAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 722 AATGCGCATCTCTCGCTAAGCTTTCATGACAGAGAAAGCTTGCATCCCTCGCTG 781
Qy 541 TTCTCAGAGTCCATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 782 TTCTCAGAGTCCATGTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 841
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 842 GTGTTCTCTCCACAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
Qy 661 GAGACACATGTGCGGAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGAGTTG 720
Db 902 GAGACCCCGTACTCCAGATTCGAGACAAAGTTCCAGAGATCGGCTCGAGCGGAGT 961
Qy 721 GGTGACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 962 GGTGACAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
Qy 781 ACTGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 1022 CCGAGCCGCTGACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
Qy 841 ATTCTCACTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 1082 ATCATGTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
Qy 901 GGGCAGGTTGTTAATCTGAGATCAAGTCCGAGTTTGAAGAGAGAGAGAGAGAGAG 960
Db 1142 GGGCAGGTTGTTAATCTGAGATCAAGTCCGAGTTTGAAGAGAGAGAGAGAGAGAG 1201
Qy 961 ATAAAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1202 ATTAAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1261
Qy 1021 GATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1262 GACGCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1321
Qy 1081 GATATTTCTTGAAGTCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1322 CACATTTCTGCGCTCCCTTCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1381
Qy 1141 TTGAAAAAGCTGAGCATACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1382 TTGG--AGGATGAGCCCTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1438
Qy 1201 AAGAGTTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1439 GAGAGTTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1498
Qy 1261 GCGCTTGTGCGCAGATTAATTTAGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1499 GCTTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558
Qy 1321 AAGACAAATATCCAGATTCAGATATCTATGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1559 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1618
Qy 1381 TCTTGCCAAATTTACAGCTGATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1619 TCTTGCCAGTTCACTGCTGATCTCATGCGCATGAGAGAGAGAGAGAGAGAGAGAG 1678
Qy 1441 ACTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1679 ACTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1738
Qy 1501 TTCACTCTCTGCTCTCACTCCGTTGTAATGATGATGATGATGATGATGATGATGATG 1560
Db 1739 TTCACTCTCTGCTCTCACTCCGTTGTAATGATGATGATGATGATGATGATGATGATG 1798
Qy 1561 AACATTTGTTTCCCTGCTGATATGAGATATATCTTCCCTTACCGAGAGAGAGAGAG 1620
Db 1799 AACATTTGTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1858
Qy 1621 AGGTTGAAGATTTCCATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1859 CGGTTGAAGATTTCCATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1918
Qy 1681 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db 1919 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1978
Qy 1741 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1979 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2038
Qy 1801 GAGTTGCTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 2039 GAGTTGCTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2098
Qy 1861 GAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 2099 GAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158
Qy 1921 TTGAGATGATATATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 2159 TTGAGATGATATATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2218
Qy 1981 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 2219 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2278
Qy 2041 GTGAGAGCAATGATCTGAGGTTTGCAGAACTTGCAGAACTTGCAGAACTTGCAGAA 2100
Db 2279 GTGAGAGCAATGATCTGAGGTTTGCAGAACTTGCAGAACTTGCAGAACTTGCAGAA 2338
Qy 2101 ATTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2339 ATTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2398
Qy 2161 GACATATCTGCTGATCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2399 GACATATCTGCTGATCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2458
```


QY 1321 AAGCAAAATATCCAGATTGAGATATCTATTGGAAAGAGCTTGAAGCAATACCATTTC 1380
DB 1764 AAGACGAAGTACCCGAGAGTACAGATATATCTGAAAGAAATTTGAGAAAGATACCACTTC 1823
QY 1381 TCTTGCCAAATTTACAGCTGATCTTTTTCAGATGAAACCATACAGATTTCATCATCCAGT 1440
DB 1824 TCTTGCCAGTTTCACTGATCTTCATCCCATGAAACCAACCGACTTATATACCAACGC 1883
QY 1441 ACTTTCAGAAATTCAGAAAGCAAGCACTGTGTCAATACGAGGCCACTGCT 1500
DB 1884 ACCTTCAGAAATTCCTGGAAGCAGAGATACAGTGGGCGATAGAGTCAATGAC 1943
QY 1501 TTCACTTCTCTGCTCTACCGTGTGTATCATGTATCATGTGTGTGATCCAAATTC 1560
DB 1944 TTCACTTCTCTGAGCTCTACCGAGTTGTTCACGGGATCGAGTCTTGACCCGAAAGTTC 2003
QY 1561 AACATTTGTTCCCGTGTGTGATATGAGATATCTTCCCTTACACCGAAAGAGAGGG 1620
DB 2004 AACATTTGTTCAACGAGTGTGACATGAGCATCTACTTGTCTTACACCGAAAGAGAGGG 2063
QY 1621 AGTTGAAGCAATTCATCTGAGATCGAAGACCTTCTTTCACCAAGTTGAGATGAA 1680
DB 2064 CGTTGAAGATCTCTCAACCTGAGATCGAAGACCTCTTTCAGGATGTTGAGAACAG 2123
QY 1681 GAACACTTATGTGTCTCAATGACCGCAACAAAGCCAAATCTGTTCAATGCGAAGCTT 1740
DB 2124 GAACACTTGTGTGTGAAAGATGAAAGAAAGCTTATATTTTTCACATGCGAAGGCTG 2183
QY 1741 GATGTGTCAAGAACTTAAACCGGACTCGTCAAGTGTGCGGCAAGAACCCAAAGTTGCT 1800
DB 2184 GACCGTGTCAAGAACTTGAACAGGGCTGTGTGATGTGTGCGAAGAACTTCAAGTTGAGG 2243
QY 1801 GAGTTGGCTAACTCGTATGTGTGATGTGTGATGAGCAAGAACTTAAAGATTGGA 1860
DB 2244 GAATCTGCGCAACTGTGTGTGTGTGAGAGTGAAGAGAAAGATTCAGAGACTTGGAA 2303
QY 1861 GAGAAAGCTGAATGAAAGAAATGTTTGAAGTGAATGCAATGCAACTTGAACGCGCAA 1920
DB 2304 GAGAGCTGTGATGAAAGAAATGTAAGACCTCATCGAAAGTCAAGCTGAATGGCCAG 2363
QY 1921 TTCAGATGATATCATCTCAATGAAACGAATCCGAATGTTGAACCTTACCAGTACTT 1980
DB 2364 TTCAGATGATATCTCTCCAGATGAAACGAGTGAAGAACTTACCGCTACATC 2423
QY 1981 TGGCAGCAAGAAAGTGCCTTTGACAGGCTGATGTGAAAGCTTGTGAATGCACTT 2040
DB 2424 TGTGACAGAAAGAGTCTTGTGATGCAACCTTGTGCACTTGTGCACTTGTGACGCTG 2483
QY 2041 GTGAGGCAATGACTTGGCGTGTGCAACATTCGCAACCTGTAAAGGTGAACAGCCGAG 2100
DB 2484 GTTGAAGGCAATGACTTGTGATGCAACCTTGTGCACTTGTGCACTTGTGACGCTGAG 2543
QY 2101 ATTATTTGCTCATGGAATCTGTGTTCAACATTTGATCTTACATGATGATCAAGCTGCT 2160
DB 2544 ATCATTTGCTCATGGAATCTGTGTTCAACATTTGATCTTACATGATGATCAAGCTGCT 2603
QY 2161 GACATACTGCTGATTTCTTTGAAAGTGTGAAAGAAAGTCAATCTCACTGGGATTAAGTC 2220
DB 2604 GAGCTTCTTTGATGATCTTCTTCAACAGTGTGCAAGATTTGACAGCTCCATCTGGGAGAGATC 2663
QY 2221 TCCCAAGAGGCTTGAACGAATAGAGAGATATACATGAAAGATTTACTCGGAGAGA 2280
DB 2664 TCAAGAGGCTGATGACAGAAATGAAAGAGATTAATGAAAGATTAATTTGAGAGG 2723
QY 2281 CTATTTGACCTTGAACAGAGTGTATGATTTGGAAGCAATTTTCAACCTTGAAGCCGT 2340
DB 2724 CTGTTGAACCTGACGCTGTGTATGCTTGTGAAGCAATGATCAATCACTTGAATCGGCGC 2783
QY 2341 GAGAGTGTGCTTACTTGAAGTGTATGCTTTAAGTACCGTAAGCTGAGTCA 2400
DB 2784 GAGAGTGTGCTTACTTGAAGTGTATGCTTTAAGTACCGTCACTGAGCACTGAGCACTG 2843
QY 2401 GTTCCATTTGCAAGAGAGTAAA 2422

DB 2844 GTTCTCCGCTGTGAGTAAA 2865
RESULT 12
US-09-938-842A-1620
; Sequence 1620, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1620
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1620
Query Match 56.6%; Score 1486.6; DB 9; Length 2427;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
QY 4 GCTGAGGTGCTCTACTCGGCTTCAAGTCTCCGTGAGCGTTGATGAGACCTTCTT 63
DB 10 GCTGAGGTATGATTAACGCGCTTCAAGCAGCAAGTGAAGCTTTGAAAGAAACCTTGT 69
QY 64 GCTGAGGAGCAAGATTTTGGCGCTTCTGAGATGAGAGGCAAGAAAGAAAT 123
DB 70 TCTGAGAGAAACGAGTCTTCTGCTTCTTCCAGGCTTGAAGCAAGGTAAGATTT 129
QY 124 CTGCAACCAATCAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAACAGAAAG 183
DB 130 TTACAACAAACAGATCATTTGCTGAATGGAAGCTTGTGCTGAACAAACCGAGAA 189
QY 184 CTGCTAATGTGCTATTTTGAAGTATGAGGCTAGTCAAGAGCGATGCTGCT 243
DB 190 CTTGAGGTGCTCTTCTTCTTGAACCTTCAATCACTCAGAGCAATGTGTGCA 249
QY 244 CCATGGTGTGCACTGCTGCTGCTGCTGCAAGGCTGTGTTGGAGTACATTAAGTGA 303
DB 250 CCATGGTGTGCTGCTGCTGCTGCTGCAAGGCTGTGTTGGAGTACATTAAGTGA 309
QY 304 GTTCAAGCCCTGTTGTTGAGAGAACTCACTGTTGCTGATCTCACTTCAAGAAAG 363
DB 310 CTGCTGATGAGTGAAGTGAATTTCACTTGAAGTGAATTTCACTTCAAGAAAG 369
QY 364 CTTGTTGAGAGTCAATGAAATCTTGTGTTGAATTTGAAATTTGAGCCCTTCAAC 423
DB 370 CTGCTGATGAGTGAAGTGAATTTCACTTGAAGTGAATTTGAGCCCTTCAAT 429
QY 424 TCATCATTTCCCGGCAACCTTCAATTTCAATTTGATGATGATGATGATGATGAT 483
DB 430 GCGTCAATCCCTGCTCAACCTTCAATTTCAATTTGATGATGATGATGATGATGAT 489
QY 484 GGTCACTTTCGAGAAATTTGCTCATGACAGAGAGAGCAACCTTGTGCTGAATTC 543
DB 490 GGTCACTTTCGAGAAATTTGCTCATGACAGAGAGAGTGTGCTTCAATGCTTAAGTTC 549
QY 544 CTGAGAGTCAATTTGCAAGAGGCAAGCAATGATGTTGAATGACAGATTTCAAGCTTG 603

Db 550 CTTCGCTTCACAGCCACGAGGCAAGAACTGATGTTGACGGAAGANTCAGAACCTC 609
Qy 604 AATGCTCTTCACATGTTTGAAGAAAACAGAGAGATCTTGGTACCTACTCTGAG 663
Db 610 AACCTCTTCACACACTTGAAGAAAACAGAGAGATCTTGGTACCTACTCTGAG 669
Qy 664 ACACCATGTCGCAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGGGGT 723
Db 670 ACACTGATGAAAGTTTGAAGCCAGTTTGAAGAGATGTTGTTGAGAGGGAATGGGA 729
Qy 724 GACACCGCAAGAGCGTCTGAGATGATCCATCTCTTTGGATCTTCTTGAAGCACT 783
Db 730 GACATGAGAGCGTCTGAGATGATGATCTCTTTTGAAGCTCTTGAAGCGCT 789
Qy 784 GATCTCTGACCTTGAAGATCTTGGAGAAATCCCATGATGTTCAATGTTGAT 843
Db 790 GATCTCTGACCTTGAAGATCTTGGAGAAATCCCATGATGTTCAATGTTGAT 849
Qy 844 CTCACTCCCAAGCACTTCTGCTCAAGCAATGTTTGGGATTCCTGCAACCGGTTGG 903
Db 850 CTCTCTCAAGATGTTTCTTGGTCAAGCAATGTTTGGTTCCTGCAACCTGTTGGA 909
Qy 904 CAGGTTGTTACATCTGATCAAGTCCGAGCTTGAAGATGAGTCTCTCCGTTA 963
Db 910 CAGGTTGTTACATCTGATCAAGTCTCTGAGATGAGTCTCTCAACGTTA 969
Qy 964 AAGCAACAGGATCAATCACTCCCTGAGATCCCATATTTACTGACCTCTCTGAT 1023
Db 970 AAGCAACAGGATCAATCACTCCCTGAGATCCCATATTTACTGACCTCTCTGAT 1029
Qy 1024 GCTGTCGCAACAGATCGGTCACAGCTTGAAGAAATGATACGCAACAGACCTCGAT 1083
Db 1030 GCGGTGAGAACTACATCGGTCAGAGCTCTGAGAGATTTGATGATGATGATGAT 1089
Qy 1084 ATTCTTGAGTACCTTTCAGAAACAGAAAGGATTTGTCAGAAATGATCTCAAGATT 1143
Db 1090 ATTCTTGAGTACCTTTCAGAAACAGAAAGGATTTGTCAGAAATGATCTCAAGATT 1149
Qy 1144 GAAAAAGTCTGCAATCTTGGAAACCTACAGAGATGTTGTCATGAAATCTCCAA 1203
Db 1150 G--AAGTCTGCAATCTTGGAAACCTACAGAGATGTTGTCATGAAATCTCCAA 1206
Qy 1204 GAGTTGCAAGGACGCGCATCTGATCATCGAAACNACAGCAGCGCATATCGTCC 1263
Db 1207 GAATTTGATGCAAGCCTGACCTTATCATTTGTAATCATGATGAAATCTTGTGCT 1266
Qy 1264 TCCTTGCTCGCAATTAATTAAGTGTCAACAGTGCACATGCTCCATGCTTGAAG 1323
Db 1267 TCTTATGCTCAAACTTGTGTCACTGATGATCATGCTCATGCTCTTGAAGAA 1326
Qy 1324 ACAAATATCCAGATTCAGATTCATTTGAAAGAGCTTGAAGAAATACATTTCT 1383
Db 1327 ACAAAGTACCGGATTCAGATTCATTTGAAAGAGCTTGAAGAAATACATTTCT 1386
Qy 1384 TGCCAAATTACAGTATCTTTTTCATGAAACATACAGATTCATCATCAGTACT 1443
Db 1387 TGCCAGTTCACTGCGGATTTTTCATGAAACATCAGATTCATCATCAGTACT 1446
Qy 1444 TTCCAGAAATTCGAGAGAGAGACATGTTGTCAATACAGAGCCACATGCTTTC 1503
Db 1447 TTCCAGAAATTCGAGAGAGAGACATGTTGTCAATACAGAGCCACATGCTTTC 1506
Qy 1504 ACTCTTCTGCTCTACCGTGTGTACATGATTCGATGTTGATCCCAATTTCAAC 1563
Db 1507 ACTCTTCTGCTCTACCGTGTGTACATGATTCGATGTTGATCCCAATTTCAAC 1566
Qy 1564 ATTGTTCCCTGCTGATATGAGATATATCTCCCTTACACCGAGAGAGCGGAGG 1623
Db 1567 ATTGTTCTCTGCTGCTGATATGAGATATATCTCCCTTACAGAGAGAGAGGTTGA 1626
Qy 1624 TTGAAGATTTTCATCTGAGATGAGAGCTTCTTTACACCAAGTTGAGATGAGAA 1683
Db 1627 TTGAAGATTTTCATCTGAGATGAGAGCTTCTCTACAGGATGTTGAGAAACAAAG 1686

Qy 1684 CACTTATGTTGCTCAATGACCGCAACAGCCAAATCTGTTCAATGCAAGGCTGAT 1743
Db 1687 CACTTATGTTGCTCAAGCAAGAAAGAGCCGATCTCTTCAATGAGGCTGAT 1746
Qy 1744 CGTCTCAAGATTTACCGGATCTCTGATGTTGCTGCGGCAAGAACCGAAGTCCGAG 1803
Db 1747 CGTCTCAAGATTTACCGGATCTCTGATGTTGCTGCGGCAAGAACCGGCTGCGAG 1806
Qy 1804 TTGCTCAACCTCGATGTTGATGTTGATAGGCGAAGAAATCTAAGATTTGGAAG 1863
Db 1807 CTAGCTTATGTTGTTGTTGTTGAGAGAGACAGAGAGAAAGTCAAGCAATGGAAG 1866
Qy 1864 AAGCTCAAAATGAGAAATGTTGATGATGATGATGATGATGATGATGATGAT 1923
Db 1867 AAGCTCAAAATGAGAAATGTTGATGATGATGATGATGATGATGATGATGAT 1926
Qy 1924 AGATGATATCATCTCAATGATGATGATGATGATGATGATGATGATGATGAT 1993
Db 1927 AGATGATATCTCTCTCAATGATGATGATGATGATGATGATGATGATGATGAT 1996
Qy 1984 GACAGAAAGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 2043
Db 1987 GACAGAAAGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 2046
Qy 2044 GAGGCAATGATCTTGGGTTTCCCAATGCACTGTAACGTTGATGATGATGAT 2103
Db 2047 GAGGCAATGATCTTGGGTTTCCCAATGCACTGTAACGTTGATGATGATGAT 2106
Qy 2104 ATTGCTCAATGAGAAATCTGTTTCAATGATGATGATGATGATGATGATGAT 2163
Db 2107 ATTGCTCAATGAGAAATCTGTTTCAATGATGATGATGATGATGATGATGAT 2166
Qy 2164 ATACTGCTGATCTTCTTGAAGAAAGTGAAGAAATGATGATGATGATGATGAT 2223
Db 2167 ACTTCTGCTGATCTTCTTGAAGAAAGTGAAGAAATGATGATGATGATGATGAT 2226
Qy 2224 CAAGGAGCTTGAAGAAATGAGAGAGATGATGATGATGATGATGATGATGAT 2283
Db 2227 AAAGGAGGCTTCAAGAGATGAGAGAGATGATGATGATGATGATGATGAT 2286
Qy 2284 TTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2343
Db 2287 TTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2346
Qy 2344 AGTGTGTTTACCTTGAATGTTTATGCTTGAATGATGATGATGATGATGAT 2403
Db 2347 GCTGCGGTTACCTTGAATGTTTATGCTTGAATGATGATGATGATGATGAT 2406
Qy 2404 CCATTGCGAGAGA 2417
Db 2407 CCTCTTGCAGAGA 2420

RESULT 13
US-09-938-842A-1620
; Sequence 1620, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111

;; PRIOR FILLING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 1620
;; LENGTH: 2427
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-938-842a-1620

Query Match 56.6%; Score 1486.6; DB 11; Length 2427;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;

Qy 4 GCTGAGCGTCTCTCATCTCGGCTCCAGTCTCCGTGAGCGTTGGATGAGACCTTCTT 63
Db 10 GCTGAACGTATGATPACCGCGCTCCAGCCAAAGCTGAGCGTTGAAACGAAACCTTGT 69
Qy 64 GCTCAGAGAAAGAGATTTTGGCCCTTGCTCAAGATCGAGGCGAAAGAAAGAAATT 123
Db 70 TCTGAGAGAAAGAGATCTTGGCTTCTCCAGGGTTGAAGCCAAAGTAAAGTATT 129
Qy 124 CTGCAACACATCAAAATTATTCTAGAGTTGAAGCTATCCCTGAAGAGACAGAAAG 183
Db 130 TTACAAACAAACAGATCATCTGCTGAATTGCAAGCTTGCCTGAACAAACCGAAGAA 189
Qy 184 CTCGCTAATGATGATTTTGAAGTATTGAAGCTAGTCAGAAAGCATCTGTGCTT 243
Db 190 CTTGAAGGTGCTCTTCTTGAACCTTCTCAAAATCCTCAGGAAGCAATTGTGTGCA 249
Qy 244 CCAATGGTTGACCTTGTCTGTCGTCAGAGCGCTGTTGGAGTACATTAAGTGAAT 303
Db 250 CCAATGGTTGCTCTAGCTGTAAGGCGCAAGCGCTGTTGGAGTACATTAAGTGAAT 309
Qy 304 GTTCACGCGCTTGTGTGAGAGAACTCACTGTGCTGATCTTCACTTCAAGAAAG 363
Db 310 CTCATGCTCTTGTGTTGAAGAACTCAACCTGCTGATTTCTTCAATTCAGAAAGAA 369
Qy 364 CTTGTGATGAGAGTTCAATGAGAACTTTGTTGAAATGATTTGAGCCCTTCAAC 423
Db 370 CTTGTGATGAGAGTTAAGATGATTAATTTCACTTGTGAGCTTGAATTCAGCCATTCAT 429
Qy 424 TCATCATTTCCCCCGCAACTTTTCAAAATCCANTGTTANAGTGGAGATTCCTAAT 483
Db 430 GCGTCAATCCCTGTCGCAACCTCCACAAATACATGGAATGAGTGTGACTTCTTAAAC 489
Qy 484 CGTCACTTTGCGCAAAATTTGTCATGACAGAGAGATGCAACCTTTGCTGAAATTC 543
Db 490 CGTCAATTTATGGCTTAAGCTCTTCCATGACAAAGAGATTTGCTTCCATTTAAAGTTC 549
Qy 544 CTCAAGTCCATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAACTTG 603
Db 550 CTTGCTCTTCAACGCCACCAAGGCGAAAGCTGATGTTGAGCGAGAAATTCAGAACTTC 609
Qy 604 AATGCTCTTCAACAGTTTGAAGAAAGCAGAGAGATCTTGTGATCCCTACCTCCAG 663
Db 610 AACACTCTGCAACACCTTGAAGAAAGCAGAAAGATCTTGAAGACCTTAAAGTCCGAA 669
Qy 664 AACACATGTGCGCAATTTGCAACACCGGTTCCAGAAATCGGTTTGAAGAGTGGGCT 723
Db 670 AACATGTATGAAGAGTTGAGGCCAAAGTTGAGAGATTTGGTCTTGAAGAGGAGATGGGA 729
Qy 724 GACACCGAGAAAGCGTCTCGAGATGATCAACTCTTTTGGATCTTTTGAAGCAACT 783
Db 730 GACATGAGAGAGCGTCTTGAACATGATGATCTTTTGGACCTTCTTGAAGCGCTT 789
Qy 784 GATCTTGGACACCTTGAAGATGTTCTTGGAGAAATCCCATGAGTTCATATGTTGTGATT 843
Db 790 GATCTTGGACACCTTGAAGATGTTCTTGGAGAGAACCAATGATGTTTAAAGTGTGATC 849
Qy 844 CTCACATCCCAAGGATATCTGCTCAAGACATGTTTGGGGTATCCGGAACACGAGTGG 903
Db 850 CTCCTCTCAACATGATTTACTTGTCTCAGACATGTTCTTGGTTACCTCGAACATGATGGA 909
Qy 904 CAGGTTGTTTATCATCTTGGATCAAGTCCGAGCTTGGAGAAATGAGATGCTCCTCGTATA 963

Db 910 CAGGTTGTTTATCATCTTGGATCAAGTCCGAGCTTGGAGATGAGATGCTTCAACGTAAT 969
Qy 964 AAGCAACAGAGATCAACATCAACCCCTGATCCCTCATTAATTAAGTATGATCTTCTGAT 1023
Db 970 AAGCAACAGAGATCAACATTAAGCAAGAGATTTCTATTTACTGACATTTCTACCTGAT 1029
Qy 1024 GCTGTGGAACCAATGCGGTCAACGACTTGAAGAAAGTATACGGAACAGAGACCTCGAT 1083
Db 1030 GCGGTAGAACTCATGCGGTGAACGCTCGAGAGAGATTAATGATTTCTGAGTACTGATGAT 1089
Qy 1084 ATTCTTGGAGTACCTTCAAGAACAGAAATTTGTTGAAATGATGATTTCTCAAGATT 1143
Db 1090 ATTCTTGGAGTACCTTCAAGAACAGAAAGGATTTGTTGCAAAATGATTTCAAGGTTT 1149
Qy 1144 GAAAGATGTGGCATCTTGAAGAACTTACAGAGAGATGTTGCTCATGAATCTCAAA 1203
Db 1150 G--AGTCTGGCATTTCTAGAGATTTACACGAGATGCTGGGTTGAGCTATCGAA 1206
Qy 1204 GAGTTGACCGGCAACCGCATTTGATATCGAAACCAAGCGGCAATATCTGCGC 1263
Db 1207 GAATTTGAATGGCAAGCTGACCTTATCATTTGTTAATCAAGTATGGAATCTTGTGCT 1266
Qy 1264 TCTTGTGCTGACATTAATTAAGTGTCAACAGTGCACATCGCCCATGCTTGGAGAG 1323
Db 1267 TCTTATTTGGCTCAAACTTGTGTCTCAGTACATGTCATGCTCTTGAAGAA 1326
Qy 1324 ACAAATATCCAGATTCAGATATCTATTTGAAGAAAGCTTGAAGCAATATCATTTCTCT 1383
Db 1327 ACBAAGTACCGGATTTGATATCTACTGGAAGAGCTTGAACAGATACATTTCTCA 1386
Qy 1384 TGCCAAATTAACGCTGATCTTTTGAAGAAACCAATCAAGATTTATCATCACAGTACT 1443
Db 1387 TGCCAGTTCACTGGGATATTTTGGCAATGAAACAACATGATTTATCATCATCAATGACT 1446
Qy 1444 TTCCAGAAATTTGAGAGAACCAAGAGACATGTTGTCAATACAGAGCCACATGCTTTC 1503
Db 1447 TTCCAAAGAAATTTGTTGAGAACCAAGAACTGTGGGAGATGATGAAGCCACACAGCTTT 1506
Qy 1504 ACTCTTCTGCTCTCAACCGGTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1563
Db 1507 ACTCTTCCGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1566
Qy 1564 ATTGTTTCCCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1623
Db 1567 ATTGTTTCCCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
Qy 1624 TTGAAGATTTTCCATCTGAGATGAGAGACCTTCTTACACCAAGTTGAGAAATGAAGAA 1683
Db 1627 TTGAAGATTTTCCATCTGAGATGAGAGACCTTCTTACACCAAGTTGAGAAATGAAGAA 1686
Qy 1684 CAATTATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1743
Db 1687 CAATTATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746
Qy 1744 CGTGTCAAGAACTTAAACCGGACTGTGTGAGTGTGTGGGCAAGAAACCAAAAGTTGCTGAG 1803
Db 1747 CGTGTCAAGAACTTGTGAGGCTTGTGAGTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGAG 1806
Qy 1804 TTGGCTAAACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863
Db 1807 CTACCTAACTTGT 1866
Qy 1864 AAGGCTGAAGTGAAGAAATGTTGAGCTGTGACAAAGTATCAACTTGAACCGGCAATTC 1923
Db 1867 AAGGCTGAAGTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1926
Qy 1924 AGATGATATCATCTCAATGAACAGATCGAAATGTTGAATCTTACCGATACATTTGC 1983
Db 1927 AGATGATATCTCTCAATGAACAGATCGGATGAAGAAAGTGTGAGCTGTACCGGATACATCTGT 1986
Qy 1984 GACAGAAAGTGTGCTTGTGACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2043

```

Db      1987 GACACCAAGGAGCTTTTGTCCAACTGCATTAATGAAAGCTTTGGTTAAGTGTGTG 2046
Qy      2044 GAGGCAATGACTGTGGCTTTGGCAACATTTGGCAACTGTAAACGGTGAACGCGAGATT 2103
Db      2047 GAGGCTAAGACTGTGGTTTACCAAGCTTTCGCACTTCAAAAGGTGTCAGCTGAGATC 2106
Qy      2104 ATTGTGCAATGGGAAATCTGTTTCAACATTTGATCTCTTACATGTGATCAAGCTGTAC 2163
Db      2107 ATTGTGCAACGGTAAATCGGGTTTCCACATTTGACCTTACATGTGATCAAGCTGTAT 2166
Qy      2164 ATACTGTCGATTTCTTTGAAAAGTGAAGAAAGATTCATCTGAGATTAAGTCTCC 2223
Db      2167 ACTGTGTCGATTTCTTCAACCAAGTGAAGAGATTCATCTCACTGGAGATGAGATCTCA 2226
Qy      2224 CAAGGAGCTTGAACCAATGAGAGAGATTAATGAGAAATTTACTCGAGAGACTA 2283
Db      2227 AAGAGAGGCTTCAAGAGATTAAGAGAAATACCTTGGCAATCTATTCACAGAGCTC 2286
Qy      2284 TTGACCCCTGACAGAGATGATGATTTCTGAAAGCATGTTTCAACCTTGAACGCGTGAG 2343
Db      2287 TTGACATTTGACTGTGTGTATGATTTCTGAAAGCATGTCTGAACTTGAACCTTGTAG 2346
Qy      2344 AGTCGTCGTAACCTTGAGATTTTATGCTCTTAAGTACCGTAACTGAGCTGAATCACTT 2403
Db      2347 GCTGCGCGTTACCTTGAATGTTCTATGATGATTAAGATTCGCGCATTTGCTCAGGCTGTT 2406
Qy      2404 CCATTGGCAGAGA 2417
Db      2407 CCTCTTGACAGA 2420

```

RESULT 14

```

US-10-425-114-9355
; Sequence 9355, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9355
; LENGTH: 2279
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700843719_FLI
US-10-425-114-9355

```

Query Match 50.7%; Score 1331; DB 13; Length 2279;

Best Local Similarity 78.8%; Pred. No. 0;

Matches 1599; Conservative 0; Mismatches 426; Indels 3; Gaps 1;

```

Qy      404 TGGATTTGAGCCCTTGAACATCATATCCCGCGCAACTCTTCAAAATCATGGTA 463
Db      1 TGGACTTTGAACCAATTCATGATGAGCTTCCCTCGCCCAACTTTAAACAAGTCAATGGAA 60
Qy      464 ATGGTGTGAGATCTCTTAATGTCACCTTTGGCAAAATTTGTCATGACAGAGAGACA 523
Db      61 ATGGTGTGAGATCTCTTCAACCGCACTTTGCGCAAACTTTCACAGACAGAGAGACT 120
Qy      524 TGCACCTTTGTGTCATTTCTTCAAGTTCATTTGTCACAGAGGCAAGAAATGATTTGA 583
Db      121 TGCACCCCACTTTTGAATTTCTTCAAGGCTTCAAGGTCACAGGAAAGACTTTGATTTGA 180
Qy      584 ATGACAGATTCAGAACTTGAATGCTTTCAACATGTTTGAAGAAAGAGAGATATC 643

```

```

Db      181 ATGACAGATTCAGAACTTGAATGCTTTCAACATGTTTGAAGAAAGCTGAGAGATATC 240
Qy      644 TTGTGACCTTCTCTTGAAGACACCATGTGCCAAATTTGAACACCGGTTCCAGAGAAATCG 703
Db      241 TGGGCAAGTGTCTCTGAAGCTCCCTTCACTGAAATTTGAACAAAGTTCCAGAGAGATTTG 300
Qy      704 GTTTGGAAGAGGTGGGGTGACACCGCAGAACCGGTGCTGAGATGATTCACATCTCTT 763
Db      301 GTTTGGAAGAGGTGGGGTGACACCGCAGAACCGGTGCTTGAATGATTCACATCTCTT 360
Qy      764 TGAATCTTCTTGAAGCAACTGATCTTGAACCTTGAAGAGTTCTTGGAGAAATCCCA 823
Db      361 TGAATCTTCTTGAAGCCCTGACCTTGAACCTTGAAGACTTCTTGGAGAAATCTCTA 420
Qy      824 TGTGTTCAATGTTGTGATTTCTACCTCCCAAGATTTCTGCTCAGACAAATGTTTGG 883
Db      421 TGTGTTCAATGTTGTGATTTCTCTCCCATGTGTACTTGGCCCAAGATATGCTTGG 480
Qy      884 GGTATCCCGACACCGGTGGGAGGTGTTTACATCTTGAATCAAGTCCGAGCTTTGAGA 943
Db      481 GATACCTGACACTGTGTGGCCAGGTGTTTACATCTTGAATCAAGTGTGCTTTGAGA 540
Qy      944 ATGAGATCTCTCTGATTAAGCAACAGGACTCAATCAACCCCTCGAATCTCATTA 1003
Db      541 ACGAGATCTCATATGCAATTAAGCAACAGGATTTGACATTTGATCTGATTTCTCATTA 600
Qy      1004 TTACTAGACTTCTCTGATGCTGTGCGAACACATGCGGTCAACGACTTGAAGAAATAT 1063
Db      601 TCACCCGCTTCTCTCCCATCAATTCGGAATCACTGTGTGGCCAACTCTTGAAGAGTGT 660
Qy      1064 ACGGAACGAGACCTCGGATTTCTGAGTACCTTCAGAACCAAGAAAGGAATGTTTC 1123
Db      661 TCGGAACGAGACCTCCCACTTCTGAGTTCCTTTGAATCTGAGAGAGAAATGTTTC 720
Qy      1124 GAAATGATCTCAAGATTTGAAAAAGCTGCGCATTACTTGGAAACCTTACA CAGAGATG 1183
Db      721 GTCAGTGATCTCAAGANTTG---AAGTCTGGCATTTTGGAAACTTACACTGAGATG 777
Qy      1184 TTGCTCATGAATCTCCAAAGATTTGACGCGCAACCGCATCTGATATCATCGGAACACNA 1243
Db      778 TTGCTCATGACCTTGCCAAAAGTTGCAAGGCAACCGCATCTGATTTGCGAAACTTCA 837
Qy      1244 GCGACGGAATATGTGCGCTCTTGTGCGACATTAATTTAGTGTACACAGTGCACA 1303
Db      838 GTGATGAAACATTTGTGCTTGTGTGCAATTAATTTAGTGTACCTGATGTAC 897
Qy      1304 TCGCCCATGCTTTGAGAGAGACAAATATCCAGATTCAGATATCTAATTTGAAGAGCTTG 1363
Db      898 TTGCTCATGACCTTGAAGAGACAAATATCCCGAATCCGACATTTACTGGAAGAAATTTGG 957
Qy      1364 AAGCAATACCATTTCTCTTGCATTAACAGCTGATCTTTTGAATGAACCATACAG 1423
Db      958 AAGAGATACCATTTCTCTTGCATTAACAGCTGATCTTTTGAATGAACCATACAG 1017
Qy      1424 ATTTCATATACAGATCTTTTGAAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1483
Db      1018 ATTTCATATACAGATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077
Qy      1484 ACGAGAGACACTGCTTCACTCTTCTGCTCTTCAACCGTGTGTCATGATGATGATG 1543
Db      1078 ATGAGTCTACAGAGCTTTTACCTTCTGAGACTTCAACCGTGTGTCATGAGAGAGAGATG 1137
Qy      1544 TGTTCATCCAAATTCACATTTGTTTCCCTGCTGTGATGAGATATATCTTCCCTT 1603
Db      1138 TCTTCATCCAAATTCACATTTGTTTCCCTGAGAGATCAACATTTTAACTTCCCTT 1197
Qy      1604 ACACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1663
Db      1198 CACCGGAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
Qy      1664 CCAAGTTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1723

```

```
Db 1258 GCTCTGTGAGAAATGAAAGACATATGCGTCTGAAGACCGCAGCAAGCCGATTAATCT 1317
Oy 1724 TCACAATGCCAAGCGTTGATCGTGTCAAGAACTTAACCGGACTCTGTCAGTGTGCGGCA 1783
Db 1318 TCACCATGCGAAGTTGACCGGTGTAAGAACATACAGAGACTCGTGAAGTGTGTCGGTA 1377
Oy 1784 AGAACCACAAAGTTGGTGAAGTTGGCTAACCTCTGTAGTGTGTAGTGTGTATAGCGCAAGG 1843
Db 1378 AGAACGCGCAACCTGGAGGAGTTGGTGAACCTTGTGTGTGTGTGCGCGAAGACGAGAGAGG 1437
Oy 1844 AATCTAAAGATTTTGGAGAGAGAGGCTGAATGAGAAATGTTTGAAGTGTGATCGACAAGT 1903
Db 1438 AGTCAGAGACTTGGAGAGAGAGGCGAGATGAGAGATGATGAGGCTGTGATCGAGACCT 1497
Oy 1904 ACAACTTGAACGCGCCCAATTCAGATGATGATCACTCAATGAGAACGAATCCGAAATGTTG 1963
Db 1498 ACAAGTTGAACGCGGCAATTCAGATGATGATCACTCAATGAGAACGATGAGAGAACGAG 1557
Oy 1964 AACTTTACCGATCACTTTGGCAGACAGAAAGTGCCTTTGTACAGCTTCGATTTGATGAG 2023
Db 1558 AGCTGTACCGTGTGATCTGCGCACACCAAGGAGCTTGTGTGACGCGCGCTATATACGAG 1617
Oy 2024 CTTTGTGATGACAGTGTGAGGAGCAATGACTTGGGTTTGGCAATTTGCAACCTGTA 2083
Db 1618 CTTTGTGATGACAGTGTGAGGAGCAATGACTTGGGTTTGGCAATTTGCAACCTGTA 1677
Oy 2084 ACGGTGACCAAGCCGAGATTAATGTCATGAGAAATCTGTTCAACTGATCTTACC 2143
Db 1678 ATGGTGTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1737
Oy 2144 ATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2203
Db 1738 ATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1797
Oy 2204 CTCACGTGGATTAAGATCTCCCAAGAGGCTTGAAGAGATGAGAGAGATTAATACGTA 2263
Db 1798 CCCACTGAGAAACATCTCAAGAGGCTGTCTCCAGCGATGAGAGAGATTAATACGTA 1857
Oy 2264 AGATTACTCGAGAGACTATTTGACCTGACAGAGATGATGATGATGATGATGATGATGAT 2323
Db 1858 AATTTTATCTCAAGAGGCTGTCTCACTCACTGATGATGATGATGATGATGATGATGATGAT 1917
Oy 2324 CCAACTTTGAAGCGCTGTGAGAGTGTGTTACCTTGAAGATGTTTATGCTCTTAAGTACC 2383
Db 1918 CTAACCTTGAAGCGCTGTGAGAGGCGCTGTATCTGAGATGATGATGATGATGATGATGAT 1977
Oy 2384 GTAAGCTGTGATCAATGATTCATTTGAGAGAGAGATTAATGAACCTG 2431
Db 1978 GCAATTTGAGCGAGTCTGTGCTTGTGCTTGTGATTAATGAGATG 2025

RESULT 15
US-10-289-757-133
; Sequence 133, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forester, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A11189, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Kelch Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; FILE OF INVENTION: grasses and methods for their use
; FILE REFERENCE: 11000.1061U
; CURRENT APPLICATION NUMBER: US/10/289,757
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/337,703
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 133
; LENGTH: 2714
; TYPE: DNA
; ORGANISM: Lolium perenne
US-10-289-757-133

Query Match 49.7%; Score 1305.4; DB 15; Length 2714;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;

Oy 16 CTCACGTGGGCGCCACAGTCTCCGAGAGGTTTGGATGAGACCTTCTGTCTACAGAAC 75
Db 57 CTGACTGTCTCCACAGTCTCCGAGAGGCTTGTGTGACCTTCTCTCCATCTTAAC 116
Oy 76 GAGATTTTGGCTTGTCTCAAGATCGAGGCAAGGAAAGAAATTTGCAACACCAT 135
Db 117 GAGCTCAATGGCCCTTTTCAAGATGATGATCAACAGGCAAGAAATGCTTCAAGGCGAC 176
Oy 136 CAATTTATCTAGAGTTTGAAGCTATCCCTGAAAGAAACAGAAAGAACTGCTTAATGCT 195
Db 177 CAGCTGTCACTGAGTTTGAAGCCTTGTGTGAGGCTGACAGAGAGA-----GATATGCA 230
Oy 196 GCATTTTGTGAAGATTTGAAGGCTAGTCAGGAGAGCATGCTGTGCTCCATGAGGTTGCA 255
Db 231 CCTTTGAAGCATTTCTCCGCTGCTCTGAGAGCAATTTGTGCTGCCCATGAGGTTGCA 290
Oy 256 CTGCTGTCTGCTCCAGGCGCTGTGTGTGAGAGTACATTAAGATGATGATCAAGCCCTT 315
Db 291 CTGCGCATCAAGGCGCTAGGCGCTGTGTGTGAGAGTACATTAAGATGATGATGATGATGAT 350
Oy 316 GTTGTGAGAACTCACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
Db 351 GCTGTGAGAGGCTGACAGATTTGAGAGTCTGAGATTCATGAGAAAGAGCTTGTGTGATGAG 410
Oy 376 AGTTCAATGAGAACTTGTGTGTGAGATTTTGAAGGCTTCACTCACTCACTCACTCACT 435
Db 411 CAGCCAGAGCAAGATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 470
Oy 436 CGCCCAACTCTTTGAAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
Db 471 GCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 530
Oy 496 GCAAAATTTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
Db 531 TCAGAGCTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590
Oy 556 TGTCAAGAGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
Db 591 AACCAAGAGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
Oy 616 CAGTTTGAAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
Db 651 TCAGCCCTTGAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 710
Oy 676 GAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 735
Db 711 GAGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 770
Oy 736 CGGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
Db 771 GGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 830
Oy 796 CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 855
Db 831 TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
Oy 856 GGAATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
Db 891 GGAATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 950
Oy 916 ATCTTGAATCAAGGCGAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 975
Db 951 ATCTTGAATCAAGGCGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1010
```


Oy 976 CTCAACATCACCCCTGAACTCTCATTAATACTAGATTCTTCTGATGCTGCGAACA 1035
 Db 1011 CTTGACATTAACCCCTTAAGATCTCTCATTTGTGCACAGGCTGTGCTGATGCTGTGAAC 1070
 Oy 1036 ACATGCGGTCAACGACTTGAGAAAGTATACGGAACAGACACTCGGATATTTCTGAGTA 1095
 Db 1071 ACATGTGCGCAGCGGTGAGAAAGTATATGGGACTGAGACACTGACATTTCTCGGT 1130
 Oy 1096 CCCCTTCAAGAACGAAAGGGAATTTGTGAAAATGGAATCTCAAGTTTGAAAAATCTGG 1155
 Db 1131 CCATTCAAACTGAAAGGGGATCTCCGTAGTGAATCTTCGTTTG--ATGTCGG 1187
 Oy 1156 CCATCTCTGGAACCTACACAGAGGATGTGCTCATGAATCTCAAAAGATGACAGGC 1215
 Db 1188 CCAATCTCGAGACATACACGAGGATTTTGCAACGAACTATAGAGAAATGACAGCC 1247
 Oy 1216 ACGCAGATCTGATCATCGAAACNACAGCGCAATATCGTGCCTCTGCTGCGCA 1275
 Db 1248 AAGCTGATCTCATTAATGGCAACTACAGTACGGTAACTGTGCGCACTGCTGCG 1307
 Oy 1276 CATTAATTAAGTGTACACAGTGCACATCGCCCATGCTTTGAGAAAGAAATATCA 1335
 Db 1308 CATAGTGGGAGTTACCCAGTGCACATGCCCATGCTTTGAGAAACAAATATACCC 1367
 Oy 1336 GATTGAGATATCTATTGGAAAGCTTGAAAGCAAAATCAATTTCTTGGCAATTTCA 1395
 Db 1368 AACTGACATATATTGGACAAATTCAGACGCAATTCATTTCTCATGCGCACTGCT 1427
 Oy 1396 GCTGATCTTTTGGCAATGACACATACAGATTTCAATCACAGTACTTTCAGGAAAT 1455
 Db 1428 GCTGACCTTATTGGCATGATACACTGATTTCAATCAACAGACATTCAGGAAAT 1487
 Oy 1456 GAGAGACGAAGGACACTGTTGCTCAATACAGAGCCACACTGCTTTCACTTCTGCT 1515
 Db 1488 GCTGAAACGAAGGACGCTGGGCGAGTATGAGTCTCAATGCTTTCACTTCTGCT 1547
 Oy 1516 CTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
 Db 1548 CTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
 Oy 1576 GGTGCTGATATGAGATATACTTCCCTTACACGGAAGAAAGCGAGGTTGAAGCATTC 1635
 Db 1608 GAGAGAGATATGACGTCTACTTCCATACACTGAACGAAAGGCTTACTGCTTTC 1667
 Oy 1636 CATCTGATGATGAAACCTTTTACACAAAGTTGAGATGAAGAACTTATGCTG 1695
 Db 1668 CACCTGAATTTGAAGACTCTCTTACAGTGAATGTTGAGAACTGGAACACAAATTTGA 1727
 Oy 1696 CTCAATGACCGCAACAAAGCAATTTCTGTCACAATGCCAAGGCTGATGCTCAAGAAC 1755
 Db 1728 TTGAAGGACAAAGAACAGCTTATCATCTTCAATGCTCGTGTGACCGTGTGAAGAAC 1787
 Oy 1756 TTAACCGGACTGTCGAGTGTGCGGCAAGAACCCAAAGTTCGAGTGGCTTAACCTC 1815
 Db 1788 ATGACAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1847
 Oy 1816 GTAGTTGAGTGTGATAGCGGAAAGAAATCTAAAGATTTGAAAGAGAGGCTGAAATG 1875
 Db 1848 GTGATGTGCTGAGCAATGCAAGAGTCAAGAGATAGGAGAGACAGGCTGAGTTTC 1907
 Oy 1876 AAGAAATGTTTGAAGCTATGACAAAGTAACTGAAACGCGCAATTCAGATGATATCA 1935
 Db 1908 AAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
 Oy 1936 TCTCAATGAAACAGAAATCGAAATGTTGAACCTTACCGATACATTTGCGACACGAAAGT 1995
 Db 1968 GCTGAGATGAACCGTGTGCGAATGCAAGATGTAACCGTACATTTGTGACACCAAGGCG 2027
 Oy 1996 GCCTTTGACAGCTGCAATGTATGAGGCTTTGATGACAGTTGTGAGGCAATGACT 2055
 Db 2028 GCATTTGTTACGCTGCAATGTATGAGCAATTTGCGTACTGTTGTGAGGCAATGACA 2087

Oy 2056 TCGGTTTGGCAACATTCGCAACCTGTACAGGTGAGACCGCCGAGATTATGTCATGGG 2115
 Db 2088 TGTGTTTGGCCCAATAGGACACATGCAATGTGGCCGAGCTGAATCATTTGTGAATGTT 2147
 Oy 2116 AAATGTGTTTCAACATTTGATTCCTTACCATGTGTATCAAGTGTGACATCTGTCAT 2175
 Db 2148 GTGCTGTGTTTGCATGATTCCTTACCAAGTGTGACAGGCTGCAAGATATCTGTTCAAC 2207
 Oy 2176 TTTCTTGAAGATGTAAAGATCCATCTCACTGGGATTAAGTCTCCCAAGAGGCTTG 2235
 Db 2208 TTTCTTGAAGATTCACCGGACATCCATCTTACGCAAAATGTGGAAGGTGGCTTG 2267
 Oy 2236 AAACGATAGAGAGAAAGTATACATGGAAGATTTACTGAGAGACTATGACCTGACA 2295
 Db 2268 AAGGAATTTATGAGAAATGACCTGGAAGCTGTACTAGAGAGGCTGATGACCTGACT 2327
 Oy 2296 GAGTGTATGATTTCTGGAACATGTTTCAACCTTGAAACCGCTGAGAGTGTGCTTAC 2355
 Db 2328 GTGTATATGTTTCTGGAAGTATGTGAGCAACCTTGAGAGGCGTGAACCTGCGTTAC 2387
 Oy 2356 CTTGAGATGTTTATGCTCTTAACTTAACTGACCTGAGCTGCTGATGATGATGATGATGAT 2415
 Db 2388 CTTGAGATGTTTCTAGCTTGAAGTACCGACGCTGCTGCTGCAATGCTTCAATGCTGCT 2447
 Oy 2416 GA 2417
 Db 2448 GA 2449

Search completed: June 5, 2004, 09:02:07
 Job time : 1545 secs

THIS PAGE BLANK (USPTO)